

THÈSE

de Docteur - Ingénieur

présentée

A LA FACULTÉ DES SCIENCES DE PARIS

par

GINO CILIGOT-TRAVAIN

pour obtenir

le grade de Docteur - Ingénieur

SUJET DE LA THÈSE

**Estimation bayésienne et identification des systèmes
gouvernés par des équations aux dérivées partielles.**

soutenue le 29 Mai 1970 devant la Commission d'Examen

MM. J. L. LIONS Président

P. A. RAVIART {
A. BENSOUSSAN } Examinateurs

THÈSE

de Docteur - Ingénieur

présentée

A LA FACULTÉ DES SCIENCES DE PARIS

par

GINO CILIGOT-TRAVAIN

pour obtenir

le grade de Docteur - Ingénieur

SUJET DE LA THÈSE

**Estimation bayésienne et identification des systèmes
gouvernés par des équations aux dérivées partielles.**

soutenue le 29 Mai 1970 devant la Commission d'Examen

MM. J. L. LIONS

Président

P. A. RAVIART

A. BENSOUSSAN

} Examineurs

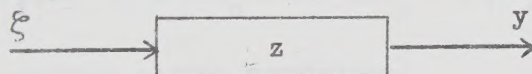


INTRODUCTION GÉNÉRALE

Le développement de la théorie du contrôle d'une part, la puissance croissante des moyens de calcul d'autre part, ont donné une nouvelle importance à la théorie des systèmes. Etant donné un système quelconque (physique, biologique, économique ou autre) le premier problème consiste à lui donner un modèle mathématique. Ce modèle dépend généralement de divers paramètres qu'il s'agit ensuite d'identifier sur le vu du fonctionnement habituel du système.

Dans ce travail, nous nous sommes intéressés aux problèmes d'identification des systèmes gouvernés par des équations aux dérivées partielles linéaires, que nous pouvons poser comme suit :

Considérons un système représenté par une boîte noire déterministe



où ξ désigne l'entrée, z l'état du système et y la sortie (ou observation). Cette sortie est une fonction linéaire de l'état perturbée par un bruit η :

$$y = Fz + \eta$$

l'état est la solution (unique) d'une équation aux dérivées partielles linéaire que nous représentons par :

$$\begin{cases} Pz = B\xi + f \\ A_j z = \xi_j \quad j = 1, \dots, N \end{cases}$$

où P est un opérateur qui dépend d'un paramètre u qu'il s'agit d'identifier " sur le vu de l'observation y .

Il existe différentes façons d'aborder ce problème, nous en présenterons essentiellement deux : la méthode bayésienne et la méthode des moindres carrés.

LA METHODE BAYESIENNE

Rappelons brièvement le principe dans le cas des systèmes de dimension finie. Soit encore u le paramètre à estimer et y l'observation, on désignera par $p(u)$ la loi de probabilité a priori sur u et $p(u|y)$ la loi a posteriori (sur le vu de l'observation y). D'après la formule de Bayes, ces lois sont reliées par :

$$p(u|y) = \frac{p(y|u) p(u)}{p(y)}$$

On appelle estimé bayésien $\hat{u}$ la valeur de u_0 qui minimise la quantité :

$$J_1(u_0) = E \left[(u_0 - u)' Q(u_0 - u) | y \right]$$

(Q étant une matrice carrée symétrique).

Par dérivation, on obtient facilement :

$$(0.1) \quad \hat{u} = E[u|y] = \int u p(u|y) du$$

Cette théorie a été étendue aux systèmes de dimension infinie par BENSOUSSAN [1]. La principale difficulté de cette extension est de nature probabiliste : il faut faire du calcul de probabilité dans les espaces de Hilbert, ce qui nécessite l'introduction des mesures cylindriques. Nous n'avons plus alors d'interprétation aussi simple qu'en dimension finie et ne pouvons reprendre le schéma précédent. Toutefois, dans le cas d'observation de type bruit blanc, on obtient un résultat analogue à (0.1) : l'estimé bayésien s'interprète comme l'espérance mathématique d'une loi de



Digitized by the Internet Archive
in 2026

probabilité a posteriori. Nous pouvons alors réitérer le processus, c'est-à-dire considérer la loi a posteriori comme nouvelle loi a priori, et ainsi, par récurrence, une loi a posteriori et un estimé à l'ordre n . La première partie de ce travail est consacrée à l'étude (théorique et numérique ⁽¹⁾) de la convergence de tels processus.

LA METHODE DES MOINDRES CARRES

La méthode des moindres carrés diffère essentiellement de la précédente par le fait qu'on n'utilise pas les propriétés stochastiques du bruit. On définit l'estimé au sens des moindres carrés comme la valeur de u qui minimise :

$$J_2(u) = \| y - \mathcal{F}_z(u) \|^2$$

($\| \cdot \|$ désigne la norme dans l'espace des observations).

Différentes variantes sont possibles pour déterminer la valeur de u optimale (algorithmes de minimisation de fonctionnelle par exemple, cf. CEA [2]). Dans le cas d'un problème évolutif (équations de type paraboliques par exemple) on peut utiliser un algorithme récursif qui présente l'avantage d'être adapté au "temps réel" : on détermine l'estimé au pas $(n + 1)$ à partir de l'estimé au pas n et de l'observation du système dans l'intervalle de temps $(t_0 + n \Delta t, t_0 + (n+1) \Delta t)$ séparant les deux pas. Nous étudierons dans la deuxième partie un algorithme proposé par SRIDHAR et PEARSON [10] et son application aux systèmes aux dérivées partielles.

Parmi les autres approches, signalons les méthodes directes, présentées brièvement en complément, et les travaux de CHAVENT [4] et EMSELLEM [6] .

(¹) Pour avoir une idée de la valeur pratique des méthodes proposées, on se reportera directement aux commentaires et conclusions accompagnant les essais numériques.

Que Monsieur LIONS trouve ici l'expression de ma très respectueuse reconnaissance pour m'avoir permis d'entreprendre ce travail.

Ma gratitude va également à Monsieur LAUDET, Directeur de l'Institut de Recherche d'Informatique et d'Automatique, pour les facilités qu'il m'a accordées dans sa réalisation.

Je tiens à remercier tout particulièrement Monsieur BENSOUSSAN, qui est à l'origine de ce travail, et n'a cessé par la suite de me prodiguer conseils et encouragements.

Je remercie Monsieur RAVIART de nous faire l'honneur de participer au jury.

Je ne puis oublier tous mes camarades de l' I.R.I.A. pour leur collaboration quotidienne, en particulier G. CHAVENT pour l'aide qu'il m'a apportée dans la connaissance des équations aux dérivées partielles.

Je remercie enfin Mademoiselle RAVET et Monsieur POMMIER pour le soin et la diligence apportés à la réalisation de ce document.

PLAN DETAILLE

<u>PREMIERE PARTIE : LA METHODE BAYESIENNE</u>	Pages
I. <u>LA TECHNIQUE BAYESIENNE</u>	8
1. Le modèle mathématique	
1.1. Equations d'état.	
1.2. Observation.	
2. Les estimateurs bayésiens	10
2.1. Etude du 1er cas.	
2.2. Approximation.	
2.3. Etude du 2ème cas : observation de type " bruit blanc ".	
II. <u>CONVERGENCE DES ESTIMATEURS BAYESIENS RECURSIFS</u>	15
3. Récursivité avec observation unique	16
3.1. Loi a posteriori à l'ordre n .	
3.2. S est de mesure μ_0 non nulle.	
3.3. S est réduit à un point.	
3.4. Application : minimum d'une fonction.	
3.5. S est réduit à un nombre fini de points.	
3.6. S est un segment de droite dans le plan.	
4. Récursivité avec une suite d'observations.	46
4.1. Lois a posteriori à l'ordre n .	
4.2. Convergence de la loi a posteriori approchée.	
4.3. Remarque sur une hypothèse.	
III. <u>APPLICATIONS NUMERIQUES</u>	54
5. Systèmes et approximations.	

5.1.	Les systèmes.	
5.2.	Approximation de l'espace des observations.	
6.	Résultats numériques.	57
6.1.	Identification d'un paramètre.	
6.2.	Identification de deux paramètres.	
6.3.	Autre exemple.	
6.4.	Identification de quatre paramètres.	

DEUXIEME PARTIE : LA METHODE DES MOINDRES CARRES

1.	Estimation au sens des moindres carrés	75
1.1.	Position du problème.	
1.2.	Conditions nécessaires d'optimalité.	
1.3.	Résolution des équations issues de 1.2.	
1.4.	Algorithme récursif.	
2.	Identification et filtrage non linéaire	80
3.	Applications numériques	81
3.1.	Le problème continu.	
3.2.	Le problème discrétisé.	
3.3.	Résultats numériques.	

COMPLEMENT : UNE METHODE DIRECTE

1. Position du problème.
2. Equation de la diffusion à une dimension.
3. Equation de la diffusion en dimension deux.

PREMIERE PARTIE

LA METHODE BAYESIENNE

L'estimation de paramètres à partir d'observations (appartenant à un espace de HILBERT) perturbée par un bruit de type " bruit blanc " nécessite l'introduction de mesures cylindriques ou de fonctionnelles aléatoires linéaires sur l'espace des observations. L'estimateur bayésien s'interprète comme l'espérance mathématique de la loi a posteriori :

$$\hat{u} = \frac{\int_{U_0} u e^{-\Psi(u)} d\mu_0}{\int_{U_0} e^{-\Psi(u)} d\mu_0}$$

(La fonction Ψ dépend des propriétés stochastiques de l'observation).

On peut définir alors l'estimateur bayésien récursif par itération des lois a posteriori :

$$\hat{u}_n = \frac{\int_{U_0} u e^{-\Psi(u)} d\mu_0}{\int_{U_0} e^{-\Psi(u)} d\mu_0}$$

Si on itère le processus avec une seule observation, l'estimateur récursif s'écrit :

$$\hat{u}_n = \frac{\int_{U_0} u e^{-n\psi(u)} d\mu_0}{\int_{U_0} e^{-n\psi(u)} d\mu_0}$$

Le comportement de $\hat{u}_n$, lorsque n croît indéfiniment, dépend de l'ensemble S des points qui rendent ψ minimale. Si S est réduit à un point, $\hat{u}_n$ converge vers ce point qui s'interprète comme l'estimateur des moindres carrés. Dans le cas contraire, la loi a posteriori à l'ordre n converge vers une loi portée par S et $\hat{u}_n$ vers un compromis des points de S .

A partir d'une suite d'observation, l'estimateur récursif devient :

$$\hat{u}_n = \frac{\int_{U_0} u e^{-n\psi_n(u)} d\mu_0}{\int_{U_0} e^{-n\psi_n(u)} d\mu_0}$$

Nous montrerons alors que l'estimateur approché $\hat{u}_{nh}$ converge vers la valeur vraie u_0 du paramètre.

Le dernier chapitre illustrera ces résultats à partir d'essais numériques.

I. LA TECHNIQUE BAYESIENNE

1. LE MODELE MATHEMATIQUE

1.1 Equations d'état

Soient $E_0, E_1, \Phi_0, \dots, \Phi_N$ des espaces de Hilbert, $P, B, A_1, \dots, A_N$ des opérateurs appartenant respectivement à $\mathcal{L}(E_0, E_1), \mathcal{L}(\Phi_0, E_1), \mathcal{L}(E_0, \Phi_1), \dots, \mathcal{L}(E_0, \Phi_N)$. On considère un système dont l'état z appartient à E_0 et vérifie les équations :

$$(1.1) \quad \begin{cases} Pz = B\xi + f \end{cases}$$

$$(1.2) \quad \begin{cases} A_j z = \xi_j & j = 1, \dots, N \end{cases}$$

$\xi \in \Phi_0, \xi_j \in \Phi_j, j = 1, \dots, N$, représentent l'entrée du système

On suppose que P et $A_j, j = 1, \dots, N$ dépendent d'un paramètre u qu'il s'agit d' "identifier". On prendra u dans U espace de Hilbert, muni d'une loi de probabilité μ (dite loi a priori) concentrée sur un sous-ensemble borné U_0 de U . Cette loi a priori résume l'information que l'on possède sur u avant toute observation.

1.2 Observation

L'état du système étant en général inaccessible à la mesure, on observe une fonction de l'état appelée sortie du système et définie par :

$$(1.3) \quad y = \mathcal{F}z$$

y appartient à un espace de Hilbert K (espace des sorties) et $\mathcal{F}$ à $\mathcal{L}(E_0, K)$. On suppose, de plus, que cette observation se fait en présence de bruit, ce qui conduit à introduire un espace d'épreuves ω muni d'une mesure m , et une fonction aléatoire réelle linéaire $Y_k(\omega, u)$:

$$(1.4) \quad \begin{cases} k \longrightarrow Y_k(\omega, u) \\ K \longrightarrow \text{Mes}\{\omega \times U, d\rho = d\mu \times d\mu ; \mathbb{R}\} \end{cases} \quad (1)$$

Ainsi, à toute sortie k du système, l'observation fait correspondre une variable aléatoire réelle sur $\omega \times U$.

On fait sur Y_k les hypothèses suivantes :

i - u étant fixé, $Y_k(., u)$ est une fonction aléatoire réelle linéaire de K dans $\text{Mes}\{\omega, d\mu; \mathbb{R}\}$ à laquelle correspond une mesure cylindrique de Gauss g_u sur K d'espérance mathématique $\bar{Y}(u)$ et d'opérateur de covariance $\Lambda(u)$.

$$(1.6) \quad \text{ii-} \begin{cases} u \longrightarrow \bar{Y}(u) \text{ est } \mu\text{-mesurable} \\ u \longrightarrow \Lambda(u)k \text{ est } \mu\text{-mesurable, } \forall k \in K \end{cases}$$

Nous distinguerons les deux cas suivants :

1.2.1 L'application (1.4) est continue.

On peut alors écrire :

$$(1.7) \quad Y_k(\omega, u) = (Y(\omega, u), k)_K, \quad \rho \text{ presque sûrement, } \forall k \in K$$

et associer à la variable aléatoire $Y(\omega, u)$ une mesure de probabilité σ sur K .

1.2.2 L'application (1.4) n'est pas continue.

On peut alors associer à $Y_k(\omega, u)$ une mesure cylindrique sur K qui ne se réduit pas à une mesure de Radon sur K . Dans ce cas, nous nous limiterons à considérer Y_k de la forme

$$(1.8) \quad Y_k(\omega, u) = (k, \bar{Y}(u))_K + \eta_k(\omega)$$

$\eta_k(\omega)$ étant une fonction aléatoire réelle linéaire à laquelle

(1) $\text{Mes}\{\Omega, dP; \mathbb{R}\}$ désigne l'ensemble des variables aléatoires réelles sur Ω muni de la mesure P .

correspond une mesure cylindrique de Gauss de moyenne nulle et d'opérateur de covariance Λ , indépendant de u , que nous supposerons inversible.

2. LES ESTIMATEURS BAYESIENS

2.1 Etude dans le 1er cas (1.2.1)

On introduit les espaces de Hilbert suivants :

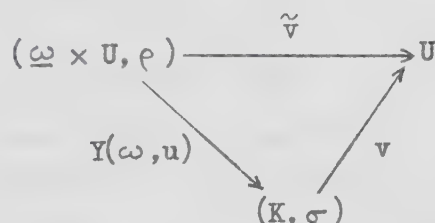
$$i - L^2(\underline{\omega} \times U, d\rho; U) = \left\{ \begin{array}{l} \text{Espace des variables aléatoires définies} \\ \text{sur } \underline{\omega} \times U, \text{ à valeur dans } U, \text{ dont le carré} \\ \text{de la norme est } \rho\text{-sommable.} \end{array} \right\}$$

$$ii - L^2(K, d\sigma; U) = \left\{ \begin{array}{l} \text{Espace des variables aléatoires définies sur} \\ K \text{ à valeur dans } U, \text{ dont le carré de la} \\ \text{norme est } \sigma\text{-sommable.} \end{array} \right\}$$

Pour simplifier, on notera $L^2(\underline{\omega} \times U, d\rho; U)$ par $L_U^2(\rho)$ et $L^2(K, d\sigma; U)$ par $L_U^2(\sigma)$.

A l'aide de $Y(\omega, u)$, on définit une application $v \rightarrow \tilde{v}$ de $L_U^2(\sigma)$ dans $L_U^2(\rho)$ par :

$$(2.1) \quad \tilde{v}(\omega, u) = v(Y(\omega, u))$$



La mesure σ étant l'image de ρ par y , cette application est une isométrie. Ainsi $\tilde{L}_U^2(\sigma)$, image de $L_U^2(\sigma)$, est un sous-espace de Hilbert de $L_U^2(\rho)$. Nous pouvons alors projeter l'application coordonnée $(\omega, u) \rightarrow u$ sur $\tilde{L}_U^2(\sigma)$, et obtenir un élément unique de $\tilde{L}_U^2(\sigma)$ auquel correspond un élément de $L_U^2(\sigma)$ que nous noterons $\hat{u}(Y)$.

Définition 1.1. L'élément $\hat{u}(Y)$ sera appelé estimateur de BAYES du paramètre u .

2.2 Approximation numérique

2.2.1 Dans la pratique $\bar{Y}(u)$ et $\Lambda(u)$ sont souvent donnés de façon implicite. On considère alors des approximations $\bar{Y}_\xi(u)$ et $\Lambda_\xi(u)$ (appartenant respectivement à K et $\mathcal{L}(K,K)$) vérifiant les propriétés :

$$(2.2) \quad i - (\bar{Y}_\xi(u), k)_K \xrightarrow[\xi \rightarrow 0]{} (\bar{Y}(u), k)_K \quad \forall k \in K, \text{ pour presque tout } u$$

$$(2.3) \quad ii - (\Lambda_\xi(u)k, l)_K \xrightarrow[\xi \rightarrow 0]{} (\Lambda(u)k, l)_K \quad \forall k, l \in K, \text{ pour presque tout } u$$

ξ appartenant à un espace d'indice de dimension finie.

2.2.2 L'espace K est généralement de dimension infinie ; pour l'approcher, on introduit une famille $\{K_h\}$ de sous-espaces vectoriels fermés de dimension finie $N(h)$, h appartenant à un espace d'indice de dimension finie. Si p_h désigne le projecteur de K dans K_h , on suppose que :

$$(2.4) \quad p_h y \longrightarrow y \text{ fortement dans } K, \quad \forall y \in K, \text{ si } h \longrightarrow 0$$

Il existe un isomorphisme I_h entre K_h et $\mathbb{R}^{N(h)}$, qui va nous permettre d'approcher Y par une variable aléatoire à valeur dans $\mathbb{R}^{N(h)}$. On pose $\tau_h = I_h \circ p_h$. Les correspondances entre les différents espaces sont représentées sur le schéma ci-dessous.

(2.5)

$$\begin{array}{ccc}
 \omega \times U & & \\
 \downarrow Y_h \xi & \searrow Y & \\
 & & K \\
 & \swarrow \tau_h & \downarrow p_h \\
 \mathbb{R}^{N(h)} & & K_h \\
 & \xleftarrow{I_h} &
 \end{array}$$

Soit $\{e_h^j\}$, $j = 1, \dots, N(h)$ la base de K_h associée à l'isomorphisme I_h et $\Lambda_{h\xi}(u)$ la matrice définie par :

$$(2.6) \quad [\Lambda_{h\xi}(u)]_{ij} = (\Lambda_\xi(u)e_h^i, e_h^j) \quad i, j = 1, \dots, N(h)$$

On fait l'hypothèse :

$$(2.7) \quad \Lambda_{h\xi}(u) \text{ est définie positive et symétrique } \forall h, \xi ; \text{ pour presque tout } u .$$

On peut alors définir une variable aléatoire $Y_{h\xi}(\omega, u)$ de $\underline{\omega} \times U$ dans $\mathbb{R}^{N(h)}$:

$$(2.8) \quad \left\{ \begin{array}{l} \text{pour } u \text{ fixé, } Y_{h\xi}(\cdot, u) \text{ suit une loi de Gauss de} \\ \text{moyenne } \tau_h \bar{Y}_\xi(u) \text{ et de matrice de covariance } \Lambda_{h\xi}(u) \end{array} \right\}$$

Définition 2.2 La variable aléatoire $Y_{h\xi}$ est appelée observation approchée du système.

La densité de $Y_{h\xi}$, pour u fixé, est donnée par :

$$(2.9) \quad \begin{aligned} f_{h\xi}^u(x_h) &= \\ &= \frac{1}{(2\pi)^{\frac{N(h)}{2}} \sqrt{\text{Det } \Lambda_{h\xi}(u)}} \exp \left[-\frac{1}{2} (x_h - \tau_h \bar{Y}_\xi(u))' \Lambda_{h\xi}^{-1}(u) (x_h - \tau_h \bar{Y}_\xi(u)) \right] \end{aligned}$$

la densité de $Y_{h\xi}(\omega, u)$ devient alors :

$$(2.10) \quad \begin{aligned} \frac{d\sigma_{h\xi}(x_h)}{dx_h} &= \\ &= \int_U \frac{d\mu(u)}{(2\pi)^{\frac{N(h)}{2}} \sqrt{\text{Det } \Lambda_{h\xi}(u)}} \exp \left[-\frac{1}{2} (x_h - \tau_h \bar{Y}_\xi(u))' \Lambda_{h\xi}^{-1}(u) (x_h - \tau_h \bar{Y}_\xi(u)) \right] \end{aligned}$$

2.2.3 On peut maintenant reprendre le schéma de 2.1. en remplaçant K par $\mathbb{R}^{N(h)}$. La projection de l'application $(\omega, u) \rightarrow u$ sur $\tilde{L}_U^2(\sigma_{h\xi})$ définit un élément $\hat{u}_{h\xi}(x_h) \in L_U^2(\sigma_{h\xi})$ vérifiant l'équation :

$$\forall v_h \in L_U^2(\sigma_{h\xi})$$

$$(2.11) \quad \int_U d\mu(u) \int_{\omega} (u, v_h(Y_{h\xi}(\omega, u)))_U d\mu(\omega) = \int_{\mathbb{R}^{N(h)}} (\hat{u}_{h\xi}(x_h), v_h(x_h)) d\sigma_{h\xi}(x_h)$$

d'où l'on tire :

$$(2.12) \quad u_{h\xi}(x_h) = \frac{\int_U \frac{d\mu(u)}{\sqrt{\text{Det} \Lambda_{h\xi}(u)}} \exp\left[-\frac{1}{2}(x_h - \tau_h \bar{Y}_{\xi}(u))' \Lambda_{h\xi}^{-1}(u) (x_h - \tau_h \bar{Y}_{\xi}(u))\right] d\mu(u)}{\int_U \frac{d\mu(u)}{\sqrt{\text{Det} \Lambda_{h\xi}(u)}} \exp\left[-\frac{1}{2}(x_h - \tau_h \bar{Y}_{\xi}(u))' \Lambda_{h\xi}^{-1}(u) (x_h - \tau_h \bar{Y}_{\xi}(u))\right] d\mu(u)}$$

Définition 2.3. On appelle estimateur de Bayes approché la variable aléatoire $\hat{u}_{h\xi}(\tau_h Y(\omega, u))$. Cette définition est justifiée par le théorème suivant :

Théorème 1.1. La suite $\hat{u}_{h\xi}(\tau_h Y)$ converge dans $L_U^2(\sigma)$ et l'on a :

$$(2.13) \quad \lim_{h \rightarrow 0} \lim_{\xi \rightarrow 0} \hat{u}_{h\xi}(\tau_h Y) = \hat{u}(Y)$$

(pour la démonstration, voir BENSOUSSAN [1])

2.3 Etude dans le 2ème cas (1.2.2). Observation de type " bruit blanc".

Lorsque l'application (1.4) n'est pas continue, on ne peut plus écrire $Y_k(\omega, u)$ sous forme de produit scalaire dans K . On suppose alors que Y_k est de la forme suivante :

$$(2.14) \quad Y_k(\omega, u) = \langle k, \bar{Y}(u) \rangle_K + \eta_k(\omega)$$

$\bar{\eta}_k(\omega)$ étant une fonction aléatoire réelle linéaire sur K à

laquelle est associée une mesure cylindrique de Gauss de type bruit blanc de moyenne nulle et d'opérateur de covariance Λ (i.e. telle que $(\Lambda\varphi, \varphi) \geq \alpha \|\varphi\|^2$, $\forall \varphi \in K$, K étant identifié à son dual).

Le schéma du paragraphe 2.1. n'est alors plus possible ; par contre, la définition de l'estimateur de Bayes approché conserve toute sa validité. Notons que la formule (2.12) se simplifie un peu, $\Lambda_{h\xi}(u) = \Lambda_h$ étant indépendant de u :

$$(2.15) \quad \hat{u}_{h\xi}(\tau_h Y_d) = \frac{\int_U d\mu(u) u \exp\left[-\frac{1}{2}(\tau_h Y_d - \tau_h \bar{Y}_\xi(u))' \Lambda_h^{-1} (\tau_h Y_d - \tau_h \bar{Y}_\xi(u))\right]}{\int_U d\mu(u) \exp\left[-\frac{1}{2}(\tau_h Y_d - \tau_h \bar{Y}_\xi(u))' \Lambda_h^{-1} (\tau_h Y_d - \tau_h \bar{Y}_\xi(u))\right]}$$

soit encore :

$$(2.16) \quad \hat{u}_{h\xi}(\tau_h Y_d) = \frac{\int_U d\mu(u) u \exp\left[\frac{1}{2}(\tau_h \bar{Y}_\xi(u))' \Lambda_{-h}^{-1} (\tau_h \bar{Y}_\xi(u)) - (\tau_h Y_d)' \Lambda_{-h}^{-1} (\tau_h \bar{Y}_\xi(u))\right]}{\int_U d\mu(u) \exp\left[\frac{1}{2}(\tau_h \bar{Y}_\xi(u))' \Lambda_{-h}^{-1} (\tau_h \bar{Y}_\xi(u)) - (\tau_h Y_d)' \Lambda_{-h}^{-1} (\tau_h \bar{Y}_\xi(u))\right]}$$

De même, on a toujours un théorème analogue au théorème 1.1.

Théorème 1.2. La suite $\hat{u}_{h\xi}(\tau_h Y_d)$ est convergente et l'on a :

$$(2.17) \quad \lim_{h \rightarrow 0} \lim_{\xi \rightarrow 0} \hat{u}_{h\xi}(\tau_h Y_d) = \hat{u}(Y_d) = \frac{\int_U d\mu(u) u \exp\left[-\frac{1}{2}(\Lambda^{-1}(Y_d - \bar{Y}(u)), Y_d - \bar{Y}(u))_K\right]}{\int_U d\mu(u) \exp\left[-\frac{1}{2}(\Lambda^{-1}(Y_d - \bar{Y}(u)), Y_d - \bar{Y}(u))_K\right]}$$

fortement dans U .

(pour la démonstration voir BENSOUSSAN [1])

On pose alors la définition suivante :

Définition 2.4. On appelle estimateur de Bayes dans ce cas la limite $\hat{u}(Y_d)$ précédente.

Nous allons étudier plus particulièrement le cas d'observation perturbée par un bruit blanc. Pour une observation (ou sortie) Y_d , l'estimateur de Bayes est donné par :

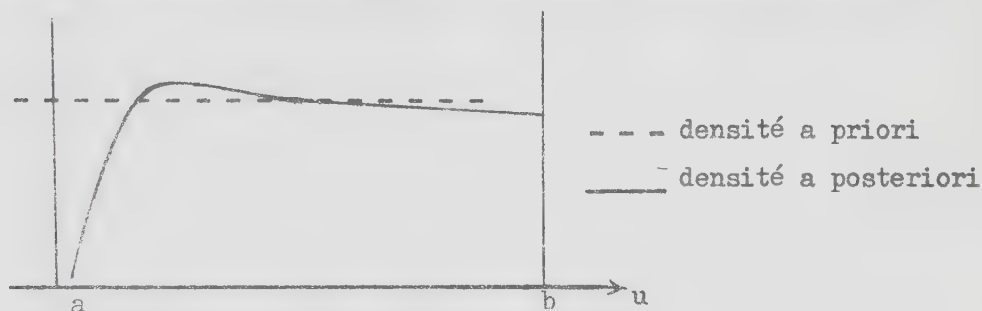
$$(2.17) \quad \hat{u}(Y_d) = \frac{\int_U d\mu_0(u) u \exp \left[-\frac{1}{2}(\Lambda^{-1}(Y_d - \bar{Y}(u)), Y_d - \bar{Y}(u))_K \right]}{\int_U d\mu_0(u) \exp \left[-\frac{1}{2}(\Lambda^{-1}(Y_d - \bar{Y}(u)), Y_d - \bar{Y}(u))_K \right]}$$

Cet estimateur s'interprète comme la moyenne de la variable aléatoire sur U_0 ayant pour densité de probabilité par rapport à la mesure a priori μ_0 :

$$(3.1) \quad \frac{d\mu_1}{d\mu_0} = \frac{\exp \left[-\frac{1}{2}(\Lambda^{-1}(Y_d - \bar{Y}(u)), Y_d - \bar{Y}(u))_K \right]}{\int_U d\mu_0(u) \exp \left[-\frac{1}{2}(\Lambda^{-1}(Y_d - \bar{Y}(u)), Y_d - \bar{Y}(u))_K \right]}$$

La loi définie par μ_1 est appelée loi a posteriori.

Les essais numériques conduisent à la difficulté suivante : la densité a posteriori diffère souvent assez peu de la densité a priori, comme le montre les courbes ci-dessous (identification d'un paramètre réel)



La moyenne est très voisine du milieu de l'intervalle (a, b) , l'estimation est peu satisfaisante. Pour remédier à cela, on utilise un procédé récursif qui consiste à prendre la loi a posteriori comme nouvelle loi a priori, à réappliquer (3.1) et ainsi de suite. Ce chapitre sera consacré à

l'étude de la convergence de la loi a posteriori à l'ordre n .

Nous envisagerons deux cas :

- (i) On conserve toujours la même observation. L'intérêt de ce cas réside dans sa simplicité pour les applications.
- (ii) On fait une nouvelle observation à chaque itération. On montre alors que l'estimateur bayésien converge, sous certaines conditions, vers la valeur réelle du paramètre à estimer.

Dans la suite, on se limitera à prendre pour U_0 un compact de $\mathbb{R}^P$, ce qui n'est pas très restrictif du point de vue des applications numériques.

3. ETUDE DE LA RECURSIVITE AVEC OBSERVATION UNIQUE.

3.1 Loi a posteriori à l'ordre n .

On désignera cette loi (ou mesure) par μ_n . On posera :

$$(3.2) \quad \psi(u) = \frac{1}{2}(\Lambda^{-1}(Y_d - \bar{Y}(u)), Y_d - \bar{Y}(u))_K$$

Lemme 3.1. La loi μ_n est absolument continue par rapport à μ_0 et définie par :

$$(3.3) \quad d\mu_n = \frac{e^{-n\psi(u)}}{\int_U e^{-n\psi(u)} d\mu_0} d\mu_0$$

Démonstration : par récurrence.

$$(3.4) \quad \text{on a :} \quad d\mu_1 = \frac{e^{-\psi(u)}}{\int_U e^{-\psi(u)} d\mu_0} d\mu_0$$

Supposons que :

$$d\mu_{n-1} = \frac{e^{-(n-1)\psi(u)}}{\int_U e^{-(n-1)\psi(u)} d\mu_0} d\mu_0$$

alors

$$d\mu_n = \frac{e^{-\psi(u)}}{\int_U e^{-\psi(u)} d\mu_{n-1}} d\mu_{n-1}$$

$$= \frac{e^{-\psi(u)} e^{-(n-1)\psi(u)} d\mu_0}{\int_U e^{-(n-1)\psi(u)} d\mu_0 \int_U e^{-\psi(u)} \frac{e^{-(n-1)\psi(u)} d\mu_0}{\int_U e^{-(n-1)\psi(u)} d\mu_0}}$$

soit
$$d\mu_n = \frac{e^{-n\psi(u)}}{\int_U e^{-n\psi(u)} d\mu_0} d\mu_0$$

On passe donc simplement de la loi a posteriori à l'ordre 1 à la loi a posteriori à l'ordre n en remplaçant ψ par $n\psi$ dans (3.4).

Orientation.

Il s'agit maintenant d'étudier le comportement de μ_n lorsque n tend vers l'infini. Soit S l'ensemble des points de U_0 qui rendent ψ minimale. En dehors de S , ψ est strictement positive ; en effet, il existe $\Lambda^{-\frac{1}{2}}$ tel que :

$$(3.5) \quad \psi(u) = \frac{1}{2}(\Lambda^{-1}(Y_d - \bar{Y}(u)), Y_d - \bar{Y}(u))_K = \frac{1}{2} \left| \Lambda^{-\frac{1}{2}} (Y_d - \bar{Y}(u)) \right|^2$$

Si m désigne le minimum de ψ sur U_0 , et si $\varphi(u) = \psi(u) - m$, nous avons :

$$(3.6) \quad d\mu_n = \frac{e^{-n\psi(u)}}{\int_U e^{-n\psi(u)} d\mu_0} d\mu_0 = \frac{e^{-n\varphi(u)}}{\int_U e^{-n\varphi(u)} d\mu_0} d\mu_0$$

La régularité de ψ intervenant seule par la suite, nous supposerons, pour simplifier, que m est nulle, c'est-à-dire que ψ s'annule sur S . La convergence de μ_n dépend du type de S ; nous allons étudier différents cas.

3.2

S est de mesure μ_o non nulle.

On désignera par χ_S la fonction caractéristique de S.

Théorème 3.1. Si : (i)- $\psi(u)$ est définie sur U_o compact de $\mathbb{R}^p$.

(ii)- le support de μ_o est contenu dans U_o

(iii)- $\mu_o(S)$ est strictement positif.

Alors μ_n tend vaguement (ou en loi) vers μ définie par :

$$(3.7) \quad d\mu = \frac{\chi_S(u)}{\mu_o(S)} d\mu_o$$

Démonstration :

Pour tout $u \in U_o$, $e^{-n\psi(u)} \longrightarrow \chi_S(u)$ si $n \longrightarrow \infty$

$$\text{si } u \in S \quad \frac{e^{-n\psi(u)}}{\int_U e^{-n\psi(u)} d\mu_o} \leq \frac{e^{-n\psi(u)}}{\mu_o(S)} \longrightarrow 0 \quad \text{si } n \longrightarrow \infty$$

$$\text{si } u \in S \quad \frac{e^{-n\psi(u)}}{\int_U e^{-n\psi(u)} d\mu_o} = \frac{1}{\int_U e^{-n\psi(u)} d\mu_o} \longrightarrow \frac{1}{\mu_o(S)} \quad \text{si } n \longrightarrow \infty$$

d'après le théorème de Lebesgue.

$$\text{donc} \quad \frac{d\mu_n}{d\mu_o} \longrightarrow \frac{\chi_S(u)}{\mu_o(S)} \quad \text{quand } n \longrightarrow \infty$$

Soit f continue, bornée sur U

$$\begin{aligned} \left| \int f(u) d\mu_n - \frac{1}{\mu_o(S)} \int_U \chi_S(u) f(u) d\mu_o \right| &\leq \int_U |f(u)| \left| \frac{d\mu_n}{d\mu_o} - \frac{\chi_S(u)}{\mu_o(S)} \right| d\mu_o \\ &\leq \sup_{u \in U_o} |f(u)| \int_U \left| \frac{d\mu_n}{d\mu_o} - \frac{\chi_S(u)}{\mu_o(S)} \right| d\mu_o \end{aligned}$$

Or $\frac{d\mu_n}{d\mu_0} \leq \frac{1}{\mu_0(S)}$ sur U_0 ; d'après le théorème de Lebesgue, le

deuxième membre tend vers zéro quand n tend vers l'infini, d'où le théorème 3.1.

3.3 S est réduit à un point $\{s\}$.

Nous dirons qu'une mesure positive m est non nulle en un point si tout ouvert contenant ce point est de mesure m strictement positive.

Théorème 3.2 : Si : (i)- $\psi(u)$ est continue sur U_0 compact de $\mathbb{R}^p$
(ii)- le support de μ_0 est contenu dans U_0
(iii)- μ_0 n'est pas nulle en s , avec $S = \{s\}$

Alors la mesure μ_n tend vaguement vers δ_s , mesure de Dirac au point s , quand n tend vers l'infini.

Démonstration :

1ère partie.

Soit V_s un ouvert quelconque contenant s ; posons $C_s = U_0 - V_s$.

Montrons que $\int_{C_s} d\mu_n \rightarrow 0$ si $n \rightarrow \infty$.

Sur C_s qui est compact, ψ est continue, strictement positive.

Il existe donc $m > 0$ tel que $\psi(u) \geq m$ sur C_s . Choisissons ε tel que $0 < \varepsilon < m$. La fonction $\psi(u)$ est nulle et continue en s , donc il existe un ouvert V_m vérifiant :

$$(3.8) \quad \begin{aligned} s &\in V_m \subset V_s \\ u \in V_m &\implies \psi(u) < \varepsilon < m \end{aligned}$$

On a alors :

$$\begin{aligned} \int_{C_s} d\mu_n &= \frac{\int_{C_s} e^{-n\psi(u)} d\mu_0}{\int_{U_0} e^{-n\psi(u)} d\mu_0} \leq \frac{\int_{C_s} e^{-n\psi(u)} d\mu_0}{\int_{V_m \cap U_0} e^{-n\psi(u)} d\mu_0} \\ &\leq \frac{\int_{C_s} e^{-nm} d\mu_0}{\int_{V_m \cap U_0} e^{-n\varepsilon} d\mu_0} = e^{-n(m-\varepsilon)} \frac{\mu_0(C_s)}{\mu_0(V_m)} \end{aligned}$$

Comme $m > \varepsilon$ et $\mu_0(V_m) > 0$ $\int_{C_s} d\mu_n \rightarrow 0$ si $n \rightarrow \infty$

2ème partie.

Montrons maintenant que μ_n tend vaguement vers δ_s .

Soit f une fonction continue bornée sur U . Il faut montrer que

$$(3.9) \quad \int_U f d\mu_n \xrightarrow{n \rightarrow \infty} f(s)$$

Considérons

$$\Delta_n = \left| \int_U f d\mu_n - f(s) \right| = \left| \int_U (f(u) - f(s)) d\mu_n \right|$$

ε étant fixé, on choisit V_s , ouvert de U contenant s , tel

$$(3.10) \quad \text{que} \quad |f(u) - f(s)| < \frac{\varepsilon}{2} \quad \text{pour} \quad u \in V_s$$

et on pose $C_s = U_0 - V_s$

$$\begin{aligned} \Delta_n &\leq \left| \int_{C_s} (f(u) - f(s)) d\mu_n \right| + \left| \int_{V_s} (f(u) - f(s)) d\mu_n \right| \\ &\leq \sup_{u \in C_s} |f(u) - f(s)| \int_{C_s} d\mu_n + \sup_{u \in V_s} |f(u) - f(s)| \end{aligned}$$

d'après la première partie, il existe N tel que

$$n > N \Rightarrow \int_{C_s} d\mu_n \leq \frac{\varepsilon}{2K} \quad (K = \sup_{u \in C_s} |f(u) - f(s)|)$$

avec (3.10), on en déduit (3.9).

Désignons par $\hat{u}_n$ l'estimateur bayésien à l'ordre n , c'est-à-dire l'espérance mathématique de la loi μ_n :

$$(3.11) \quad \hat{u}_n = \int_U u d\mu_n(u) = \frac{\int_U u e^{-n\psi(u)} d\mu_0}{\int_U e^{-n\psi(u)} d\mu_0}$$

Corollaire 3.2.1. Sous les hypothèses du théorème 3.2 $\hat{u}_n$ converge vers s lorsque n tend vers l'infini.

Démonstration. Soit U un ouvert contenant U_0 . D'après le lemme d'Urysohn, il existe une fonction g continue à support compact (donc bornée) vérifiant :

$$\left\{ \begin{array}{l} g(u) = 1 \text{ pour } u \in U_0 \\ g(u) = 0 \text{ pour } u \notin V \\ 0 \leq g(u) \leq 1 \text{ pour } u \in C_V U_0 \end{array} \right.$$

Dans la deuxième partie de la démonstration du théorème 3.2, il suffit alors de prendre $f(u) = u g(u)$ pour obtenir le corollaire, μ_n étant nulle en dehors de U_0 .

Le point s peut s'interpréter également comme un estimateur du sens des moindres carrés. (Cela découle de (3.5)). Le corollaire 3.2.1. dit alors que l'estimateur bayésien à l'ordre n converge vers cet estimateur des moindres carrés lorsque n croît indéfiniment.

3.4. Application : minimum d'une fonction

Quittons un moment l'estimation bayésienne pour considérer une fonction réelle φ quelconque définie sur U_0 .

Le théorème précédent nous permet de construire une suite qui converge vers l'abscisse du minimum de φ , s'il est unique.

Corollaire 3.2.2.

Si φ , continue sur U_0 compact de $\mathbb{R}^p$, possède un minimum unique sur U_0 en s , si l'on pose :

$$(3.12) \quad s_n = \frac{\int_{U_0} e^{-n \varphi(u)} du}{\int_{U_0} e^{-n \varphi(u)} du}$$

alors la suite s_n converge vers s lorsque n tend vers l'infini.

(C'est le corollaire 3.2.1 avec μ_0 uniforme sur U_0).

Dans la voie des applications numériques, on peut énoncer un théorème qui tient compte de l'approximation de φ par une fonction en escalier φ_h .

Etant donné un vecteur $h = (h_1, h_2, \dots, h_p)$ de $\mathbb{R}^p$, considérons le

réseau formé par l'ensemble des points a_h^k de la forme :

$$(3.13) \quad a_h^k = \sum_{i=1}^P k_i h_i$$

$k_1, k_2, \dots, k_p$, parcourant $\mathbb{Z}$. On posera :

$$(3.14) \quad W_h^k = \left\{ \text{l'hyperpavé centré en } a_h^k \text{ de côtés } h_1, \dots, h_p \right\}$$

$$(3.15) \quad U_h = \left\{ a_h^k : U_0 \cap W_h^k \neq \emptyset \right\}$$

$$(3.16) \quad \varphi_h(u) = \sum_{a_h^k \in U_h} \varphi(a_h^k) \chi_{W_h^k}(u)$$

$$(3.17) \quad d\mu_{nh} = \frac{e^{-n\varphi_h(u)} du}{\int_{U_0} e^{-n\varphi_h(u)} du}$$

$$(3.18) \quad \hat{u}_{nh} = \int_{U_0} u d\mu_{nh}$$

Théorème 3.3. Si φ , continue sur U_0 compact de $\mathbb{R}^p$, possède un minimum unique sur U_0 en s , alors μ_{nh} tend vaguement vers δ_s , mesure de Dirac au point s , et $\hat{u}_{nh}$ converge vers s lorsque h et $\frac{1}{n}$ tendent vers zéro.

Démonstration :

Comme précédemment, posons $m = \inf_{U_0} \varphi$, $\psi(u) = \varphi(u) - m$ et

$$\psi_h(u) = \varphi_h(u) - m.$$

Pour $u \in W_h^k$ on a :

$$\begin{aligned}
|\psi_h(u) - \psi(u)| &\leq \sup_{u \in W_h^k} |\psi(a_h^k) - \psi(u)| \\
&\leq \sup_{u, v \in W_h^k} |\psi(v) - \psi(u)| \\
&\leq \sup_{\substack{u \in W_h^k \\ t \in W_h^0}} |\psi(u+t) - \psi(u)|
\end{aligned}$$

d'où, pour tout $u \in U_0$:

$$|\psi_h(u) - \psi(u)| \leq \sup_{u \in U_0, t \in W_h^0} |\psi(u+t) - \psi(u)|$$

Or ψ est uniformément continue sur le compact U_0 , donc, pour ε_1 fixé, il existe H_1 tel que $|h| \leq H_1$ entraîne

$$\sup_{u \in U_0, t \in W_h^0} |\psi(u+t) - \psi(u)| \leq \varepsilon_1 \quad \text{d'où} \quad |\psi_h(u) - \psi(u)| \leq \varepsilon_1$$

pour tout $u \in U_0$.

Soit V_s un ouvert contenant s , posons $C_s = U_0 - V_s$. Sur C_s , ψ admet un minimum α strictement positif. Choisissons $\varepsilon_1 \leq \frac{\alpha}{4}$, il existe H_1 tel que $|h| \leq H_1$ entraîne $|\psi_h(u) - \psi(u)| \leq \varepsilon_1$ pour tout $u \in U_0$, d'où $\psi_h(u) \geq \alpha - \varepsilon_1$ sur C_s .

Choisissons ε_2 vérifiant $0 < \varepsilon_2 < \alpha - 3\varepsilon_1$. ψ est continue en s , donc il existe un ouvert $V'_s \subset V_s$ ($s \in V'_s$) tel que $\psi(u) \leq \varepsilon_2$ sur V'_s , d'où $\psi_h(u) \leq \varepsilon_2 + \varepsilon_1 < \alpha - \varepsilon_1$ sur V'_s .

Alors :

$$\begin{aligned}
\int_{C_s} d\mu_{nh} &= \frac{\int_{C_s} e^{-n\varphi_h(u)} du}{\int_{U_0} e^{-n\varphi_h(u)} du} \\
&= \frac{\int_{C_s} e^{-n\psi_h(u)} du}{\int_{U_0} e^{-n\psi_h(u)} du} \\
&\leq \frac{\int_{C_s} e^{-n\psi_h(u)} du}{\int_{V'_s} e^{-n\psi_h(u)} du} \\
&\leq e^{-n(\alpha - \varepsilon_2 - 2\varepsilon_1)} \frac{\int_{C_s} du}{\int_{V'_s} du}.
\end{aligned}$$

Or $\alpha - \varepsilon_2 - 2\varepsilon_1 > 0$ d'où $\int_{C_s} d\mu_{nh} \rightarrow 0$ quand h et $\frac{1}{n}$

tendent vers zéro.

Il suffit ensuite de reprendre la deuxième partie de la démonstration du théorème 3.2 pour montrer que μ_{nh} tend vers δ_s .

3.5

S est réduit à un nombre fini de points.

3.5.1. On a vu, au paragraphe 3.3 que μ_n converge vers δ_s lorsque S est réduit à un point s. Il est assez naturel de penser que μ_n converge vers une combinaison linéaire de mesures de Dirac lorsque S se compose d'un nombre fini de points. C'est ce que nous allons

démontrer dans ce paragraphe, avec des hypothèses plus restrictives que celles du théorème 3.2. toutefois. Pour simplifier un peu, nous supposerons d'abord que S est composé de deux points.

Théorème 3.4 : Si les hypothèses suivantes sont vérifiées :

H.1. S est réduit à deux points : $S = \{s_1, s_2\} \subset U_0$

H.2.1. Ψ est continue sur U_0 compact de $\mathbb{R}^p$.

H.2.2. Ψ est de classe C^3 au voisinage de s_1 et s_2 .

H.2.3. les matrices $\Psi''(s_1)$ et $\Psi''(s_2)$ sont inversibles.

H.3.1. μ_0 est absolument continue par rapport à la mesure de Lebesgue, soit $d\mu_0(u) = \rho(u)du$

H.3.2. la densité ρ est continue en s_1 , en s_2 et vérifie

$$\rho(s_1) > 0, \quad \rho(s_2) > 0$$

Alors μ_n tend vers une combinaison linéaire de mesures de Dirac, plus précisément :

$$\mu_n \xrightarrow{n \rightarrow \infty} c_1 \delta_{s_1} + c_2 \delta_{s_2}$$

avec

$$\begin{cases} \frac{1}{c_1} = 1 + \frac{\rho(s_2)}{\rho(s_1)} \sqrt{\frac{\text{Det } \Psi''(s_1)}{\text{Det } \Psi''(s_2)}} \\ c_2 = 1 - c_1 \end{cases}$$

Démonstration :

Nous utiliserons souvent la propriété suivante :

Lemme 3.4.1.

Si : (i) - Ψ est continue sur U_0 compact de $\mathbb{R}^p$

(ii) - la mesure μ_0 , concentrée sur U_0 , est non nulle au voisinage de S .

Alors, pour tout ouvert V_S contenant S , on a :

$$\int_{U_0 - V_S} d\mu_n \longrightarrow 0 \quad \text{quand } n \longrightarrow \infty$$

(la démonstration de ce lemme est analogue à la première partie de la démonstration du théorème 3.2).

1er PAS : Lemme 3.4.2. Si, pour tous voisinages ouverts V_1 de s_1 et V_2 de s_2 , on a :

$$(3.4.1) \quad \int_{V_1} d\mu_n \xrightarrow{n \rightarrow \infty} c_1$$

$$(3.4.2) \quad \int_{V_2} d\mu_n \xrightarrow{n \rightarrow \infty} c_2$$

$$\text{Alors} \quad \mu_n \xrightarrow{n \rightarrow \infty} c_1 \delta_{s_1} + c_2 \delta_{s_2}$$

Démonstration du lemme :

Soit f continue, bornée sur U :

$$|f(u)| \leq M \quad \text{pour } u \in U$$

Il faut montrer que :

$$\int_{U_0} f d\mu_n \xrightarrow{n \rightarrow \infty} c_1 f(s_1) + c_2 f(s_2)$$

Pour ε fixé, il existe deux voisinages ouverts V_1 de s_1 et V_2 de s_2 , tels que :

$$V_1 \cap V_2 = \emptyset$$

$$(3.4.3) \quad \forall u \in V_1 \quad |f(u) - f(s_1)| \leq \frac{\varepsilon}{5}$$

$$(3.4.4) \quad \forall u \in V_2 \quad |f(u) - f(s_2)| \leq \frac{\varepsilon}{5}$$

On pose $C = U_0 - (V_1 \cup V_2)$. Alors :

$$\begin{aligned}
 \left| \int_{U_0} f d\mu_n - c_1 f(s_1) - c_2 f(s_2) \right| &= \left| \int_C f d\mu_n + \int_{V_1} f d\mu_n + \int_{V_2} f d\mu_n \right. \\
 &\quad \left. - \int_{V_1} f(s_1) d\mu_n + f(s_1) \int_{V_1} d\mu_n - c_1 f(s_1) - \int_{V_2} f(s_2) d\mu_n + f(s_2) \int_{V_2} d\mu_n - c_2 f(s_2) \right| \\
 &\leq M \int_C d\mu_n + \int_{V_1} |f(u) - f(s_1)| d\mu_n \\
 &\quad + \int_{V_2} |f(u) - f(s_2)| d\mu_n + M \left| \int_{V_1} d\mu_n - c_1 \right| + M \left| \int_{V_2} d\mu_n - c_2 \right|
 \end{aligned}$$

D'après le lemme 3.4.1, il existe N_1 tel que

$$n > N_1 \implies \int_C d\mu_n \leq \frac{\varepsilon}{5M}$$

(3.4.1) et (3.4.2) entraînent :

$$\text{il existe } N_2, N_3 \text{ tels que } \begin{cases} n > N_2 \implies \left| \int_{V_1} d\mu_n - c_1 \right| \leq \frac{\varepsilon}{5M} \\ n > N_3 \implies \left| \int_{V_2} d\mu_n - c_2 \right| \leq \frac{\varepsilon}{5M} \end{cases}$$

donc, pour $n \geq \max(N_1, N_2, N_3)$ on a :

$$(3.4.5) \quad M \int_C d\mu_n + M \left| \int_{V_1} d\mu_n - c_1 \right| + M \left| \int_{V_2} d\mu_n - c_2 \right| \leq \frac{3\varepsilon}{5}$$

$$\text{d'autre part : } \int_{U_0} d\mu_n = 1$$

Alors (3.4.3) et (3.4.4) entraînent :

$$(3.4.6) \int_{V_1} |f(u) - f(s_1)| d\mu_n + \int_{V_2} |f(u) - f(s_2)| d\mu_n \leq \frac{2\varepsilon}{5}$$

De (3.4.5) et (3.4.6) découle le lemme.

2ème PAS : Il suffit en fait de montrer que $\int_{V_1} d\mu_n$ tend vers C_1

pour un voisinage V_1 particulier.

En effet, soient V_1' et V_1'' deux voisinages ouverts de s_1 , posons

$$V_1 = V_1' \cap V_1''.$$

$$\text{On a : } \int_{V_1'} d\mu_n = \int_{V_1} d\mu_n + \int_{V_1' - V_1} d\mu_n$$

$$\text{D'après le lemme 3.4.1. } \int_{V_1' - V_1} d\mu_n \rightarrow 0 \text{ si } n \rightarrow \infty$$

$$\text{d'où } \lim_{n \rightarrow \infty} \int_{V_1} d\mu_n = \lim_{n \rightarrow \infty} \int_{V_1'} d\mu_n = \lim_{n \rightarrow \infty} \int_{V_1''} d\mu_n$$

si une de ces trois limites existe.

3ème PAS : Etant donnés V_1 et V_2 voisinages ouverts disjoints de s_1 et s_2 , les deux propositions suivantes sont équivalentes :

$$(i) - \int_{V_1} d\mu_n \xrightarrow{n \rightarrow \infty} c_1 \text{ et } \int_{V_2} d\mu_n \xrightarrow{n \rightarrow \infty} c_2, \text{ avec } c_1 + c_2 = 1$$

$$(ii) - \frac{\int_{V_1} e^{-n\psi} d\mu_0}{\int_{V_2} e^{-n\psi} d\mu_0} \xrightarrow{n \rightarrow \infty} l \in [0, \infty], \text{ avec } l = \frac{c_1}{1 - c_1}$$

En effet :

$$\frac{\int_{V_1} d\mu_n}{\int_{V_2} d\mu_n} = \frac{\int_{V_1} e^{-n\psi} d\mu_0}{\int_{V_2} e^{-n\psi} d\mu_0}$$

d'où l'implication : (i) entraîne (ii).

Inversement, on sait que $\int_{V_1} d\mu_n + \int_{V_2} d\mu_n$ tend vers 1 quand n tend vers l'infini.

Comme

$\int_{V_1} d\mu_n$ est strictement positif, pour tout n , on a :

$$\begin{aligned} \int_{V_1} d\mu_n + \int_{V_2} d\mu_n &= \int_{V_1} d\mu_n \left(1 + \frac{\int_{V_2} d\mu_n}{\int_{V_1} d\mu_n} \right) \\ &= \int_{V_1} d\mu_n \left(1 + \frac{\int_{V_1} e^{-n\psi} d\mu_0}{\int_{V_2} e^{-n\psi} d\mu_0} \right) \end{aligned}$$

d'où $\int_{V_1} d\mu_n \rightarrow \frac{1}{1+l}$ quand $n \rightarrow \infty$

De même :

$$\int_{V_2} d\mu_n \rightarrow \frac{l}{1+l} \quad \text{quand } n \rightarrow \infty$$

On a ainsi l'implication : (ii) entraîne (i).

Il nous reste à montrer que $\frac{\int_{V_1} e^{-n\psi} d\mu_0}{\int_{V_2} e^{-n\psi} d\mu_0}$ converge, lorsque

n tend vers l'infini et à calculer la limite ℓ .

L'hypothèse H.2.2. nous permet d'écrire un développement limité de ψ au voisinage de s_1 :

$$(3.4.7) \quad \psi(s_1 + h) = \psi(s_1) + \psi'(s_1) \cdot h + \frac{1}{2} \psi''(s_1)(h, h) + o(\|h\|^2)$$

$$(3.4.8) \quad \text{avec } o(\|h\|^2) \leq K_1 \|h\|^3$$

De même au voisinage de s_2 , on a :

$$(3.4.9) \quad \psi(s_2 + k) = \psi(s_2) + \psi'(s_2) \cdot k + \frac{1}{2} \psi''(s_2)(k, k) + o(\|k\|^2)$$

$$(3.4.10) \quad \text{avec } o(\|k\|^2) \leq K_2 \|k\|^3$$

Ces développements limités se réduisent en fait aux deux derniers termes, car ψ est nulle et minimale en s_1 et s_2 . Nous allons montrer

$$\text{que } \frac{\int_{V_1} e^{-n\psi} d\mu_0}{\int_{V_2} e^{-n\psi} d\mu_0} \quad \text{et} \quad \frac{\int_{V_1} e^{-\frac{n}{2} \psi''(s_1)(u-s_1, u-s_1)} d\mu_0}{\int_{V_2} e^{-\frac{n}{2} \psi''(s_2)(u-s_2, u-s_2)} d\mu_0}$$

ont même limite quand n tend vers l'infini.

4ème PAS : Evaluons la limite de

$$I_n = (\sqrt{n})^p \int_{V_1} \exp \left[-\frac{n}{2} \psi''(s_1)(u-s_1, u-s_1) \right] d\mu_0$$

Par translation, on obtient, si $h = u - s_1$:

$$I_n = (\sqrt{n})^p \int_{V_0} \exp \left[-\frac{n}{2} \psi''(s_1)(h, h) \right] \rho(h + s_1) dh$$

(V_0 désigne le translaté de V_1)

On a vu que la limite ℓ est indépendante du voisinage ouvert choisi, on prendra alors pour V_0 une boule $B(\lambda)$ centrée à l'origine, de rayon λ . Posons $y = \sqrt{n} h$, il vient :

$$I_n = \int_{B(\sqrt{n}\lambda)} \exp\left[-\frac{1}{2} \Psi''(s_1)(y,y)\right] \rho\left(\frac{y}{\sqrt{n}} + s_1\right) dy$$

Avec

$$\rho'_n(y) = \begin{cases} \rho\left(\frac{y}{\sqrt{n}} + s_1\right) & \text{si } y \in B(\sqrt{n}\lambda) \\ 0 & \text{sinon} \end{cases}$$

l'intégrale devient :

$$I_n = \int_{\mathbb{R}^p} \exp\left[-\frac{1}{2} \Psi''(s_1)(y,y)\right] \rho'_n(y) dy$$

La matrice $\Psi''(s_1)$ est carrée, symétrique (réelle) ; il existe alors une matrice orthogonale (unitaire) T telle que $T \Psi''(s_1) T^{-1}$ soit diagonale. Soit D_1 cette matrice ; nous désignerons par λ_i^1 , $i = 1, \dots, p$, les valeurs propres de $\Psi''(s_1)$. D'après H.2.3,

toutes ces valeurs propres sont strictement positives (Ψ étant minimale en s_1). Effectuons un dernier changement de variable, posons $z = Ty$. On obtient :

$$\begin{aligned} I_n &= \int_{\mathbb{R}^p} \exp\left[-\frac{1}{2} z' D_1 z\right] \rho'_n(T^{-1}z) dz \\ &= \int_{\mathbb{R}^p} \exp\left[-\frac{1}{2} \sum_{i=1}^p \lambda_i^1 z_i^2\right] \rho'_n(T^{-1}z) dz \end{aligned}$$

Quand $n \rightarrow \infty$, $\rho'_n(T^{-1}z) \rightarrow \rho(s_1)$

De plus, ρ'_n est bornée (car ρ est continue en s_1 , donc bornée au voisinage de s_1) ; alors d'après le théorème de Lebesgue :

$$(3.4.11) \quad I_n \xrightarrow{n \rightarrow \infty} \rho(s_1) \int_{\mathbb{R}^p} \exp\left[-\frac{1}{2} \sum_{i=1}^p \lambda_i^1 z_i^2\right] dz = \rho(s_1) \sqrt{\frac{(2\pi)^p}{\text{Det } \psi''(s_1)}}$$

On montre de même que :

$$(3.4.12) \quad \int_{V_2} \exp\left[-\frac{n}{2} \psi''(s_2)(u - s_2, u - s_2)\right] d\mu_0 \xrightarrow{n \rightarrow \infty} \rho(s_2) \frac{(2\pi)^p}{\text{Det } \psi''(s_2)}$$

5ème PAS : Montrons que $(\sqrt{n})^p \int_{V_1} e^{-n\psi} d\mu_0$ a même limite que

$$(\sqrt{n})^p \int_{V_1} \exp\left[-\frac{n}{2} \psi''(s_1)(u - s_1, u - s_1)\right] d\mu_0.$$

Soit

$$\Delta_n = (\sqrt{n})^p \left| \int_{V_1} (\exp[-n\psi(u)] - \exp\left[-\frac{n}{2} \psi''(s_1)(u - s_1, u - s_1)\right]) d\mu_0 \right|$$

D'après (3.4.7), en posant $h = u - s_1$ et en prenant $B(\lambda)$ pour translaté de V_1 , on obtient :

$$\Delta_n = (\sqrt{n})^p \left| \int_{B(\lambda)} \exp\left[-\frac{n}{2} \psi''(s_1)(h, h)\right] (\exp[-n O(\|h\|^2)] - 1) \rho(h + s_1) dh \right|$$

Par le même changement de variable que précédemment ($z = Th$), il vient :

$$\Delta_n = (\sqrt{n})^p \left| \int_{B(\lambda)} e^{-\frac{n}{2} z' D_1 z} (e^{-n O(\|z\|^2)} - 1) \rho(T^{-1} z + s_1) dz \right|$$

$$\text{Si } O(\|z\|^2) \leq 0, \quad \left| e^{-\frac{n}{2} z' D_1 z} (e^{-n O(\|z\|^2)} - 1) \right| \leq e^{-\frac{n}{2} z' D_1 z} - n O(\|z\|^2)$$

$$\text{Or } \frac{O(\|z\|^2)}{\|z\|^2} \rightarrow 0 \text{ quand } z \rightarrow 0$$

choisissons ε inférieur à $\frac{1}{2} \min_{i=1, \dots, p} (\lambda_i^1)$, $\varepsilon > 0$. A condition de

prendre λ suffisamment petit, on a :

$$O(\|z\|^2) \geq -\varepsilon \|z\|^2 = -z' \varepsilon I z$$

$$\text{et } e^{-\frac{n}{2} z' D_1 z - n O(\|z\|^2)} \leq e^{-\frac{n}{2} z' (D_1 - 2\varepsilon I) z}$$

$$\text{Si } O(\|z\|^2) \leq 0 \quad \left| e^{-n O(\|z\|^2)} - 1 \right| \leq 1 \quad \text{d'où}$$

$$\begin{aligned} \left| e^{-\frac{n}{2} z' D_1 z - n O(\|z\|^2)} - 1 \right| &\leq e^{-\frac{n}{2} z' D_1 z} \\ &\leq e^{-\frac{n}{2} z' (D_1 - 2\varepsilon I) z} \end{aligned}$$

Alors, si l'on prend sur chaque axe $|z_i| \geq \sqrt{\frac{\log n}{\beta_i n}} (\beta_i = \lambda_i^1 - 2\varepsilon)$, on a :

$$e^{-\frac{n}{2} z' (D_1 - 2\varepsilon I) z} = \prod_{i=1}^p e^{-n \beta_i z_i^2} \leq \frac{1}{(n)^p}$$

Comme $\sqrt{\frac{\log n}{\beta_i n}}$ tend vers zéro quand n tend vers l'infini, on peut

décomposer $B(\lambda)$ en deux parties :

$$B(\lambda) = H_n \cup C_n$$

H_n désigne l'hyperpavé défini par $|z_i| \leq \sqrt{\frac{\log n}{\beta_i n}}$, $i = 1, \dots, p$,

et C_n le complémentaire dans $B(\lambda)$.

On a ainsi :

$$\begin{aligned} \Delta_n \leq (\sqrt{n})^p &\left\{ \int_{C_n} e^{-\frac{n}{2} z' (D_1 - 2\varepsilon I) z} \rho(T^{-1} z + s_1) dz + \right. \\ &\left. + \int_{H_n} \left| e^{-n O(\|z\|^2)} - 1 \right| \rho(T^{-1} z + s_1) dz \right\} \end{aligned}$$

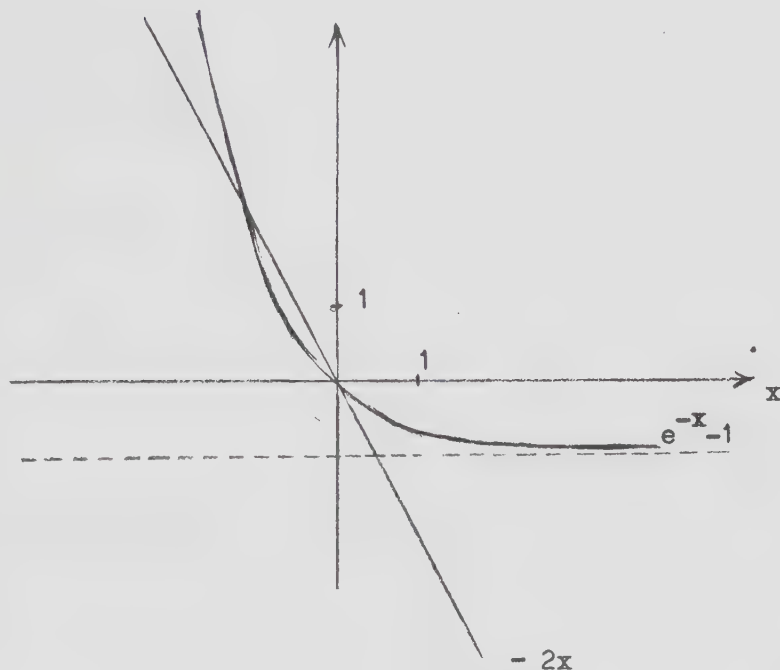
On vient de montrer que :

$$\int_{C_n} e^{-\frac{n}{2} z' (D_1 - 2\varepsilon I) z} \rho(T^{-1} z + s_1) dz \leq \frac{1}{(n)^p}$$

Il reste à majorer l'intégrale sur H_n .

D'une manière générale, pour $|x| \leq 1$ on a $|e^{-x} - 1| \leq 2|x|$

Il suffit de tracer les courbes $e^{-x} - 1$ et $-2x$ pour s'en convaincre.



Dans H_n on a :

$$\begin{aligned}
 |n O(\|z\|^2)| &\leq n K_1 |z|^3 \\
 &\leq n K_1 |z|^2 (|z_1| + |z_2| + \dots + |z_p|) \\
 &\leq n K_1 \frac{\text{Log } n}{n} \left(\sum_{i=1}^p \frac{1}{\beta_i} \right) p \cdot \max_{i=1, \dots, p} |z_i|
 \end{aligned}$$

$$\leq K_1 \left(\sum_{i=1}^p \frac{1}{\beta_i} \right) p \cdot \log n \sqrt{\frac{\log n}{n \beta_m}}$$

$$\text{avec } \beta_m = \min_{i=1, \dots, p} \beta_i$$

$$\text{Or } \log n \sqrt{\frac{\log n}{n \beta_m}} \longrightarrow 0 \quad \text{si } n \longrightarrow \infty$$

donc, dans H_n :

$$\left| e^{-n O(\|z\|^2)} - 1 \right| \leq 2 \left| n O(\|z\|^2) \right|$$

$$\leq 2 K_1 p \sum_{i=1}^p \frac{1}{\beta_i} \sqrt{\frac{(\log n)^3}{n \beta_m}}$$

$$\text{Ainsi } \int_{H_n} \left| e^{-n O(\|z\|^2)} - 1 \right| \rho(T^{-1}z + s_1) dz$$

$$\leq 2 K_1 p \sum_{i=1}^p \frac{1}{\beta_i} \sqrt{\frac{(\log n)^3}{n \beta_m}} \int_{h_n} \rho(T^{-1}z + s_1) dz$$

La densité ρ est continue en s_1 donc bornée au voisinage de s_1 :

$$\rho(T^{-1}z + s_1) \leq K \rho \quad \text{pour } z \in H_n \text{ et } n \text{ suffisamment grand}$$

d'où

$$\int_{H_n} \rho(T^{-1}z + s_1) dz \leq K \rho \left(\sqrt{\frac{\log n}{n}} \right)^p \prod_{i=1}^p \sqrt{\frac{1}{\beta_i}}$$

On peut alors majorer Δ_n :

$$\Delta_n \leq \left(\frac{1}{\sqrt{n}} \right)^p + 2 K_1 K_\rho \left(\sum_{i=1}^p \frac{1}{\beta_i} \right) \left(\prod_{i=1}^p \sqrt{\frac{1}{\beta_i}} \sqrt{\frac{(\log n)^3 + 2p}{n \beta_m}} \right)$$

Ainsi Δ_n tend vers zéro, quand n croît indéfiniment, d'où le 5ème pas.

Les pas 4 et 5 nous donnent alors :

$$\lim_{n \rightarrow \infty} (\sqrt{n})^p \int_{V_1} e^{-n\psi} d\mu_0 = \rho(s_1) \sqrt{\frac{(2\pi)^p}{\det \psi''(s_1)}}$$

On a un résultat analogue sur V_2 , d'où :

$$\lim_{n \rightarrow \infty} \frac{\int_{V_1} e^{-n\psi} d\mu_0}{\int_{V_2} e^{-n\psi} d\mu_0} = \frac{\rho(s_1)}{\rho(s_2)} \sqrt{\frac{\det \psi''(s_2)}{\det \psi''(s_1)}}$$

Le théorème 3.4 est ainsi démontré.

Ce théorème se généralise sans difficulté au cas d'un nombre fini de points. Nous pouvons énoncer le corollaire :

Corollaire 3.4.1. Sous les hypothèses suivantes :

$$H'.1. \quad S = \{s_1, s_2, \dots, s_m\}, \quad S \subset U_0$$

$$H'.2.1. \quad \psi \text{ est continue sur } U_0 \text{ compact de } \mathbb{R}^p$$

$$H'.2.2. \quad \psi \text{ est de classe } C^3 \text{ au voisinage de } s_1, s_2, \dots, s_m.$$

$$H'.2.3. \quad \text{les matrices } \psi''(s_i), i = 1, \dots, m, \text{ sont inversibles.}$$

$$H.3.1. \quad d\mu_0 = \rho(u) du$$

H.3.2 la densité ρ est continue en $s_1, s_2, \dots, s_m$ et :

$$\rho(s_i) > 0 \quad i = 1, \dots, m.$$

On a :

$$\mu_n \xrightarrow{n \rightarrow \infty} c_1 \delta_{s_1} + c_2 \delta_{s_2} + \dots + c_m \delta_{s_m}$$

$$\text{avec } \frac{1}{c_i} = 1 + \frac{\rho(s_i)}{\sqrt{\text{Det } \Psi''(s_i)}} \cdot \frac{1}{\sum_{j=1}^m \frac{\rho(s_j)}{\sqrt{\text{Det } \Psi''(s_j)}}}$$

(il suffit de généraliser les pas 1 et 3 de la démonstration précédente, les pas 2, 4 et 5 restant inchangés).

3.5.2. Vérifions sur un exemple que les hypothèses H.2.1 et H.2.2 ne sont pas incompatibles avec les problèmes étudiés numériquement.

Considérons le système défini par les équations :

$$(I) \begin{cases} \frac{\partial y}{\partial t} - \frac{\partial}{\partial x} (g(x, u) \frac{\partial y}{\partial x}) = f \text{ pour } (x, t) \in Q =]0, 1[\times]0, T[\\ y(0, t) = y(1, t) = 0 \\ y(x, 0) = y_0(x) \end{cases}$$

Si : (i) $f \in L^2(Q)$

(ii) $y_0 \in H_0^1(]0, 1[)$

(iii) $g(x, u)$ est de classe C^1 sur $]0, 1[$ en x .

alors la solution y de (I) appartient à

$$H^{2,1}(Q) = L^2((0, T); H^2(]0, 1[)) \quad .$$

Si la dimension de l'espace est inférieure ou égale à 3, pour toute fonction h de $H^2(]0, 1[)$ il existe une fonction $\tilde{h}$, égale

presque partout à h , qui est continue sur $]0,1[$.

On peut alors définir l'observation z :

pour presque tout t de $[0,T]$, $z(t,u) = y(a,t;u) + (\text{bruit blanc}), a \in]0,1[$.

Ainsi $\bar{Y}(u) = y(a,.;u)$

L'application trace $y \longrightarrow y(a,.;u)$ est linéaire continue de $H^{2,1}(Q)$ dans $L^2(0,T)$ (cf. LIONS-MAGENES [8]). D'autre part, l'application $a \longrightarrow y$ (y solution de (I) où a remplace g) est de classe C^3 de $C^1(]0,1[)$ dans $H^{2,1}(Q)$ (cf. CHAVENT [4]).

Si nous supposons enfin que g est de classe C^3 de V (ouvert de $\mathbb{R}^p$ contenant U_0) dans $C^1(]0,1[)$, ce qui est toujours le cas dans la pratique, nous obtenons les hypothèses H.2.1 et H.2.2.

3.6

Cas particulier d'un ensemble S non dénombrable de mesure μ_0 -nulle:
 S est un segment de droite dans le plan.

Le cas général d'un ensemble S non dénombrable de mesure μ_0 -nulle, est difficile à étudier, étant donné la grande variété possible de tels ensembles. Nous nous limiterons à un cas particulier rencontré dans la pratique. On prendra :

$$(3.19) \quad U_0 = [-a, +a] \times [-b, +b] \subset \mathbb{R}^2$$

(a et b étant deux réels positifs)

$$(3.20) \quad S = [0] \times [-b, +b] \subset U_0$$

De (3.20), on tire facilement que $\psi''(0, u_2)$, si elle existe, est de la forme :

$$(3.21) \quad \psi''(0, u_2) = \begin{pmatrix} 2\alpha(u_2) & 0 \\ 0 & 0 \end{pmatrix} \quad \forall u_2 \in [-b, +b]$$

On peut alors énoncer :

Théorème 3.5 : Si :

H.1. U_0 et S sont donnés par (3.19) et (3.20)

H.2.1. ψ est continue sur U_0

H.2.2. de plus, ψ est de classe C^3 sur un ouvert contenant S

H.2.2. la fonction $\alpha(u_2)$, définie en (3.21), est continue sur $[-b, +b]$ et vérifie :

$$0 < m \leq \alpha(u_2) \leq M \quad \forall u_2 \in [-b, +b]$$

H.3. la mesure μ_0 est absolument continue par rapport à la mesure de Lebesgue ($d\mu_0(u) = \rho(u) du$), ρ étant supposée continue, non constamment nulle sur S .

Alors μ_n , mesure a posteriori à l'ordre n , converge vers $\delta_{S,\lambda}$, mesure concentrée sur S définie par

$$\langle f, \delta_{S,\lambda} \rangle = \int_{-b}^{+b} f(0,y) \lambda(y) dy \quad \forall f \text{ continue, bornée avec :}$$

$$(3.22) \quad \lambda(y) = \frac{\rho(0,y)}{\sqrt{\alpha(y)}} \cdot \frac{1}{\int_{-b}^b \frac{\rho(0,y)}{\sqrt{\alpha(y)}} dy}$$

Démonstration :

1er PAS : Limite de $\sqrt{n} \iint_{V_S} e^{-n\alpha(u_2)n_1^2} \rho(u_1, u_2) du_1 du_2$, quand $n \rightarrow \infty$

(V_S étant un voisinage de S de la forme $[-d, +d] \times [-b, +b]$)

Posons

$$\begin{aligned}
 I_n &= \sqrt{n} \iint_{V_S} e^{-n\alpha(u_2)u_1^2} \rho(u_1, u_2) du_1 du_2 = \\
 &= \sqrt{n} \int_{-b}^{+b} du_2 \int_{-d}^{+d} e^{-n\alpha(u_2)u_1^2} \rho(u_1, u_2) du_1 .
 \end{aligned}$$

Pour tout u_2 on a :

$$\begin{aligned}
 \sqrt{n} \int_{-d}^{+d} e^{-n\alpha(u_2)u_1^2} \rho(u_1, u_2) du_1 &= \\
 &= \int_{-\sqrt{nd}}^{\sqrt{nd}} e^{-\alpha(u_2)y^2} \rho\left(\frac{y}{\sqrt{n}}, u_2\right) dy \\
 &= \int_{-\infty}^{+\infty} e^{-\alpha(u_2)y^2} \rho'_n(y, u_2) dy
 \end{aligned}$$

$$\text{avec } \rho'_n(y, u_2) = \begin{cases} \rho\left(\frac{y}{\sqrt{n}}, u_2\right) & \text{si } |y| < \sqrt{n} d \\ 0 & \text{sinon} \end{cases}$$

Pour tout y , $\rho'_n(y, u_2)$ converge vers $\rho(0, u_2)$ si n tend vers l'infini. De plus ρ'_n est bornée (car ρ , continue sur S , est bornée sur V_S) ; alors :

$$I_n \xrightarrow[n \rightarrow \infty]{} \int_{-b}^{+b} du_2 \int_{-\infty}^{+\infty} e^{-\alpha(u_2)y^2} \rho(0, u_2) dy = \sqrt{\pi} \int_{-b}^{+b} \frac{\rho(0, u_2)}{\sqrt{\alpha(u_2)}} du_2$$

On montre de même que, pour toute fonction f continue, bornée sur U :

$$\sqrt{n} \iint_{V_S} f(u_1, u_2) \rho(u_1, u_2) e^{-n\alpha(u_2)u_1^2} du_1 du_2 \xrightarrow{n \rightarrow \infty} \sqrt{\pi} \int_{-b}^{+b} \frac{f(0, u_2) \rho(0, u_2) du_2}{\sqrt{\alpha(u_2)}}$$

2ème PAS : Posons

$$\Delta_n = \sqrt{n} \left| \iint_{V_S} (e^{-n\psi(u_1, u_2)} - e^{-n\alpha(u_2)u_1^2}) \rho(u_1, u_2) du_1 du_2 \right|$$

A condition de prendre V_S (voisinage de S dans U_0) assez petit, on a :

$$\underline{\Delta_n \xrightarrow{n \rightarrow \infty} 0}$$

Considérons un domaine V_S de la forme : $[-d, +d] \times [-b, +b]$, $d > 0$.

$$\Delta_n = \sqrt{n} \left| \int_{-b}^{+b} du_2 \int_{-d}^{+d} (e^{-n\psi(u_1, u_2)} - e^{-n\alpha(u_2)u_1^2}) \rho(u_1, u_2) du_1 \right|$$

Fixons u_2 ; on peut écrire :

$$\psi(u_1, u_2) = \alpha(u_2) u_1^2 + o(|u_1|^2) \quad (o(.) \text{ dépend de } u_2) .$$

La fonction ψ étant de classe C^3 sur un voisinage ouvert de S (dans U), à condition de prendre d assez petit. $\|\psi'''\|$ est borné sur V_S , d'où :

$$|o(|u_1|^2)| \leq K|u_1|^3 \quad (K \text{ ne dépend plus de } u_2)$$

Si d est assez petit, on a : $K|u_1| \leq \frac{m}{2}$.

Ainsi :

$$\begin{aligned} \text{si } o(|u_1|^2) \leq 0 \quad & \left| e^{-n\psi(u_1, u_2)} - e^{-n\alpha(u_2)u_1^2} \right| \leq e^{-n\psi(u_1, u_2)} \\ & \leq e^{-n\frac{m}{2}u_1^2} \end{aligned}$$

$$\begin{aligned} \text{si } o(|u_1|^2) \geq 0 \quad & \left| e^{-n\psi(u_1, u_2)} - e^{-n\alpha(u_2)u_1^2} \right| \leq e^{-n\alpha(u_2)u_1^2} \\ & \leq e^{-n\frac{m}{2}u_1^2} \end{aligned}$$

On décompose alors l'intervalle $[-d, +d]$ en deux parties H_n et

C_n :

$$(i) \quad H_n = \left\{ u_1 : u_1^2 \leq \frac{2 \log n}{n \cdot m} \right\}$$

$$(ii) \quad C_n = \left\{ \text{complémentaire de } H_n \text{ dans } [-d, +d] \right\}$$

(en effet pour n assez grand, H_n est contenu dans $[-d, +d]$)

D'après la majoration précédente, on a :

$$\left| \frac{-n\psi(u_1, u_2)}{e} - \frac{-n\alpha(u_2) u_1^2}{e} \right| \leq \frac{1}{n} \text{ sur } C_n$$

Majorons cette quantité sur H_n :

$$\left| n O(|u_1|^2) \right| \leq nK |u_1|^3 \leq nK \left(\frac{2 \log n}{n \cdot m} \right)^{3/2} \xrightarrow{\text{si } n \rightarrow \infty} 0$$

D'où :

$$\begin{aligned} \left| \frac{-n\alpha(u_2) u_1^2}{e} - \frac{n O(|u_1|^2)}{e} \right| &\leq \left| \frac{-n O(|u_1|^2)}{e} - 1 \right| \leq \left| \frac{-n O(|u_1|^2)}{e} - 1 \right| \\ &\leq 2n |O(|u_1|^2)| \\ &\leq 2n K \cdot \left(\frac{2 \log n}{n \cdot m} \right)^{3/2} \end{aligned}$$

Alors :

$$\begin{aligned} &\int_{-d}^{+d} \left| \frac{-n\psi(u_1, u_2)}{e} - \frac{-n\alpha(u_2) u_1^2}{e} \right| \rho(u_1, u_2) du_1 \\ &\leq \frac{1}{n} \int_{C_n} \rho(u_1, u_2) du_1 + 2n K \left(\frac{2 \log n}{n \cdot m} \right)^{3/2} \int_{H_n} \rho(u_1, u_2) du_1 \end{aligned}$$

ρ étant continue sur S vérifie :

$$|\rho(u_1, u_2)| \leq K' \text{ sur } V_S$$

Ainsi

$$\int_{-d}^{+d} \left| e^{-n\psi(u_1, u_2)} - e^{-n\alpha(u_2)u_1^2} \right| \rho(u_1, u_2) du_1$$

$$\leq \frac{2dK}{n} + 16 KK' \frac{1}{n} \left(\frac{\log n}{m} \right)^2$$

D'où $\Delta_n \rightarrow 0$ si $n \rightarrow \infty$

On a de même, pour toute fonction f continue bornée sur U :

$$\sqrt{n} \left| \iint_{V_S} (e^{-n\psi(u_1, u_2)} - e^{-n\alpha(u_2)u_1^2}) f(u_1, u_2) \rho(u_1, u_2) du_1 du_2 \right| \xrightarrow[n \rightarrow \infty]{} 0$$

3ème PAS : On peut démontrer maintenant le théorème 3.5 .

Posons $\lambda(y) = \frac{P(0, y)}{\sqrt{\alpha(y)}} \cdot \frac{1}{\int_{-b}^{+b} \frac{P(0, y)}{\sqrt{\alpha(y)}} dy}$

Soit f une fonction continue, bornée sur U . On a :

$$\left| \iint_{U_0} f(u_1, u_2) d\mu_n(u) - \int_{-b}^{+b} f(0, y) \lambda(y) dy \right|$$

$$\leq \left| \iint_{U_0 - V_S} f d\mu_n \right| + \left| \iint_{V_S} f d\mu_n - \int_{-b}^{+b} f(0, y) \lambda(y) dy \right|$$

(V_S étant choisi suffisamment petit pour que la proposition du 2ème pas soit valable).

D'après le lemme 3.4.1 :

$$\iint_{U_0 - V_S} f d\mu_n \rightarrow 0 \text{ si } n \rightarrow \infty$$

D'autre part :

$$\iint_{V_S} f \, d\mu_n = \frac{\sqrt{n} \iint_{V_S} f e^{-n\psi} \rho \, du}{\sqrt{n} \iint_{U_0} e^{-n\psi} \rho \, du}$$

les pas 1 et 2 nous donnent :

$$\sqrt{n} \iint_{V_S} f e^{-n\psi} \rho \, du \xrightarrow{n \rightarrow \infty} \sqrt{\pi} \int_{-b}^{+b} f(0, u_2) \frac{\rho(0, u_2)}{\sqrt{\alpha(u_2)}} \, du_2$$

$$\sqrt{n} \iint_{U_0} e^{-n\psi} \rho \, du = \sqrt{n} \iint_{U_0 - V_S} e^{-n\psi} \rho \, du + \sqrt{n} \iint_{V_S} e^{-n\psi} \rho \, du$$

$$\text{et } \sqrt{n} \iint_{V_S} e^{-n\psi} \rho \, du \xrightarrow{n \rightarrow \infty} \sqrt{\pi} \int_{-b}^{+b} \frac{\rho(0, u_2)}{\sqrt{\alpha(u_2)}} \, du_2$$

sur $U_0 - V_S$, il existe $\ell > 0$ tel que $\psi \geq \ell$, soit $e^{-n\psi} \leq e^{-n\ell}$

$$\text{alors } \sqrt{n} \iint_{U_0 - V_S} e^{-n\psi} \rho \, du \leq \sqrt{n} e^{-n\ell} \iint_{U_0 - V_S} \rho \, du \rightarrow 0 \text{ si } n \rightarrow \infty$$

Nous avons donc :

$$\iint_{V_S} f e^{-n\psi} d\mu_n \xrightarrow{n \rightarrow \infty} \int_{-b}^{+b} f(0, y) \lambda(y) \, dy$$

le théorème est ainsi démontré.

4. ETUDE DE LA RECURSIVITE AVEC UNE SUITE D'OBSERVATIONS

4.1 Lois a posteriori à l'ordre n .

Considérons une suite d'observations $\{Y_{dn}, n \in \mathbb{N}\}$. Le passage de la loi a priori à la loi a posteriori est donné, à l'ordre n , par :

$$\frac{d\mu_n}{d\mu_{n-1}} = \frac{\exp\left[-\frac{1}{2}(\Lambda^{-1}(Y_{dn} - \bar{Y}(u)), Y_{dn} - \bar{Y}(u))_K\right]}{\int_{U_0} d\mu_{n-1}(u) \exp\left[-\frac{1}{2}(\Lambda^{-1}(Y_{dn} - \bar{Y}(u)), Y_{dn} - \bar{Y}(u))_K\right]}$$

ou encore, après simplification :

$$(4.1) \quad \frac{d\mu_n}{d\mu_{n-1}} = \frac{\exp\left[-\frac{1}{2}(\Lambda^{-1}\bar{Y}(u), \bar{Y}(u))_K + (\Lambda^{-1}Y_{dn}, \bar{Y}(u))_K\right]}{\int_{U_0} d\mu_{n-1}(u) \exp\left[-\frac{1}{2}(\Lambda^{-1}\bar{Y}(u), \bar{Y}(u))_K + (\Lambda^{-1}Y_{dn}, \bar{Y}(u))_K\right]}$$

Contrairement au cas précédent, cette transformation dépend de l'indice n (par Y_{dn}). Nous pouvons cependant énoncer un lemme assez voisin du lemme 3.1. Posons :

$$(4.2) \quad \psi_n(u) = \frac{1}{2}(\Lambda^{-1}\bar{Y}(u), \bar{Y}(u))_K - (\Lambda^{-1}(\frac{Y_{d1} + Y_{d2} + \dots + Y_{dn}}{n}), \bar{Y}(u))_K$$

Lemme 4.1 : La loi a posteriori μ_n est absolument continue par rapport à μ_0 et vérifie :

$$(4.3) \quad d\mu_n = \frac{e^{-n\psi_n(u)}}{\int_{U_0} d\mu_0(u) e^{-n\psi_n(u)}} d\mu_0(u)$$

Démonstration : par récurrence.

pour $n = 1$, on a bien :

$$\frac{d\mu_1}{d\mu_0} = \frac{e^{-\psi_1(u)}}{\int_{U_0} d\mu_0 e^{-\psi_1(u)}}$$

Supposons que :

$$\frac{d\mu_{n-1}}{d\mu_0} = \frac{e^{-(n-1)\psi_{n-1}(u)}}{\int_{U_0} d\mu_0 e^{-(n-1)\psi_{n-1}(u)}}$$

Alors

$$\begin{aligned} \mu_n &= \frac{\exp\left[-\frac{1}{2}(\Lambda^{-1}\bar{Y}(u), \bar{Y}(u))_K + (\Lambda^{-1}Y_{dn}, \bar{Y}(u))_K\right] \exp\left[-(n-1)\psi_{n-1}(u)\right] d\mu_0}{\int_{U_0} \exp\left[-\frac{1}{2}(\Lambda^{-1}\bar{Y}(u), \bar{Y}(u))_K + (\Lambda^{-1}\bar{Y}_{dn}, \bar{Y}(u))_K\right] \frac{\exp\left[-(n-1)\psi_{n-1}(u)\right] d\mu_0}{\int_{U_0} \exp\left[-(n-1)\psi_{n-1}\right] d\mu_0} \int_{U_0} \exp\left[-(n-1)\psi_{n-1}\right] d\mu_0} \\ &= \frac{\exp\left[-n\psi_n(u)\right] d\mu_0}{\int_{U_0} d\mu_0 \exp\left[-n\psi_n(u)\right]} \end{aligned}$$

La difficulté de l'étude du comportement à l'infini de μ_n provient de la moyenne $\frac{Y_{d1} + Y_{d2} + \dots + Y_{dn}}{n}$ qui n'a pas d'interprétation probabiliste simple. On se contentera d'étudier la convergence de la loi a posteriori approchée μ_{nh} définie par :

$$(4.4) \quad \left\{ \begin{aligned} d\mu_{nh} &= \frac{e^{-n\psi_{nh}(u)} d\mu_0}{\int_{U_0} e^{-n\psi_{nh}(u)} d\mu_0} \\ \text{avec } \psi_{nh}(u) &= \frac{1}{2} (\tau_h \bar{Y}(u))' \Lambda_h^{-1} (\tau_h \bar{Y}(u)) - \left(\frac{X_1 + X_2 + \dots + X_n}{n} \right)' \Lambda_h^{-1} (\tau_h \bar{Y}(u)) \end{aligned} \right.$$

$X_1, X_2, \dots, X_n$ étant des variables aléatoires indépendantes à valeur dans $\mathbb{R}^{N(h)}$, de moyenne $\tau_h \bar{Y}(u_0)$ et de matrice de covariance Λ_h , et u_0 la valeur vraie du paramètre à identifier. (τ_h, Λ_h, K_h et $N(h)$ sont définis en 2.2.2).

On notera encore :

$$(4.5) \quad \psi_h(u) = \frac{1}{2} (\tau_h \bar{Y}(u))' \Lambda_h^{-1} (\tau_h \bar{Y}(u)) - (\tau_h \bar{Y}(u_0))' \Lambda_h^{-1} (\tau_h \bar{Y}(u))$$

Montrons que $\psi_h(u)$ est minimale pour $u = u_0$.

On peut écrire ψ_h sous la forme :

$$\psi_h(u) = \frac{1}{2} (\tau_h \bar{Y}(u_0) - \tau_h \bar{Y}(u))' \Lambda_h^{-1} (\tau_h \bar{Y}(u_0) - \tau_h \bar{Y}(u)) - \frac{1}{2} (\tau_h \bar{Y}(u_0))' \Lambda_h^{-1} (\tau_h \bar{Y}(u_0))$$

Or Λ_h^{-1} est autoadjoint, positif, dont il existe $\Lambda_h^{-\frac{1}{2}}$ autoadjoint, positif,

tel que
$$\varphi' \Lambda_h^{-1} \varphi = \left| \Lambda_h^{-\frac{1}{2}} \varphi \right|^2 \quad \forall \varphi \in \mathbb{R}^{N(h)}$$

d'où :

$$\psi_h(u) = \frac{1}{2} \left| \Lambda_h^{-\frac{1}{2}} (\tau_h \bar{Y}(u_0) - \tau_h \bar{Y}(u)) \right|^2 - \frac{1}{2} \left| \Lambda_h^{-\frac{1}{2}} (\tau_h \bar{Y}(u_0)) \right|^2$$

Il est clair alors que $\psi_h(u) \geq \psi_h(u_0)$, l'égalité ayant lieu pour $u = u_0$ au moins. Dans toute la suite, nous supposons que seul u_0 rend ψ_h minimale.

4.2

Convergence de la loi a posteriori approchée.

Nous avons vu au paragraphe 3.3 que l'estimateur bayésien à l'ordre n (avec observation unique) converge vers le minimum de ψ . Dans le cas d'une suite d'observations intervient la loi des grands nombres qui va nous permettre de montrer que l'estimateur bayésien à l'ordre n converge vers la valeur vraie du paramètre à estimer.

Lemme 4.2 Si $\bar{Y}(u)$ est borné sur U_0 ($\|\bar{Y}(u)\|_K \leq M, \forall u \in U_0$), pour tout h la fonction $\psi_{nh}(u)$ converge vers $\psi_h(u)$, uniformément en u sur U_0 , pour presque toute suite d'observations.

Démonstration.

Evaluons la différence $\Delta_n(u) = |\psi_{nh}(u) - \psi_h(u)|$

$$\Delta_n(u) = \left| \left(\tau_h \bar{Y}(u_0) - \frac{X_1 + \dots + X_n}{n} \right)' \Lambda_h^{-1} (\tau_h \bar{Y}(u)) \right|$$

l'application définie par Λ_h^{-1} est bilinéaire, continue donc

$$\Delta_n(u) \leq c \left| \tau_h \bar{Y}(u) \right| \left| \tau_h \bar{Y}(u_0) - \frac{X_1 + \dots + X_n}{n} \right|$$

Soit
$$\Delta_n(u) \leq c.M. \left| \tau_h \bar{Y}(u_0) - \frac{X_1 + \dots + X_n}{n} \right|$$

D'après la loi des grands nombres, pour presque toute suite d'observations, on a :

$$\left| \tau_h \bar{Y}(u_0) - \frac{X_1 + X_2 + \dots + X_n}{n} \right| \longrightarrow 0 \quad \text{si } n \longrightarrow \infty$$

d'où le lemme.

On peut alors établir la convergence de μ_{nh} vers δ_{u_0} .

Théorème 4.1. Sous les hypothèses suivantes :

H.1 La fonction $\psi_h(u)$ est continue sur U_0 compact de $\mathbb{R}^p$;

H.2 Seul u_0 (valeur vraie du paramètre) rend ψ_h minimale;

H.3 $\bar{Y}(u)$ est borné sur U_0 .

H.4 la loi a priori initiale μ_0 n'est pas nulle au voisinage de u_0 ;

pour tout h et presque toute suite d'observations, la loi a posteriori à l'ordre n , μ_{nh} , converge (vaguement ou en loi) vers la mesure de Dirac au point u_0 , δ_{u_0} , quand n croît indéfiniment.

Démonstration

Montrons qu'en dehors de tout voisinage contenant u_0 , $\int_C d\mu_{nh} \xrightarrow[n \rightarrow \infty]{} 0$.

Soit V un voisinage ouvert de u_0 ; posons $C = U_0 - V$.

La fonction ψ_h est continue sur C , donc il existe m tel que :

$$\psi_h(u) \geq m > m_0 = \psi_h(u_0) \quad \text{pour } u \in C$$

Posons $\varepsilon_1 = \frac{m - m_0}{4}$; ψ_h est continue en u_0 , donc il existe un voisinage V_{u_0} tel que :

$$\psi_h(u) \leq m_0 + \varepsilon_1 \quad \text{pour } u \in V_{u_0}$$

D'après le lemme 4.2, il existe N_1 tel que :

$$n \geq N_1 \implies |\psi_{nh}(u) - \psi_h(u)| \leq \varepsilon_1 \quad \text{pour } u \in U_0$$

Ainsi, pour $n > N_1$, on a :

$$(i) \quad \psi_{nh}(u) \geq m - \varepsilon_1 = m_0 + 3\varepsilon_1 \quad \text{sur } C$$

$$(ii) \quad \psi_{nh}(u) \geq m_0 + 2\varepsilon_1 \quad \text{sur } V_{u_0}$$

On en déduit :

$$\begin{aligned}
\int_C d\mu_{nh}(u) &= \frac{\int_C e^{-n\psi_{nh}} d\mu_0}{\int_{U_0} e^{-n\psi_{nh}} d\mu_0} \\
&\leq \frac{\int_C e^{-n\psi_{nh}} d\mu_0}{\int_{V_{u_0}} e^{-n\psi_{nh}} d\mu_0} \\
&\leq \frac{e^{-n(m_0 + 3\varepsilon_1)} \int_C d\mu_0}{e^{-n(m_0 + 2\varepsilon_1)} \int_{V_{u_0}} d\mu_0} = e^{-n\varepsilon_1} \frac{\mu_0(C)}{\mu_0(V_{u_0})}
\end{aligned}$$

La mesure μ_0 étant non nulle au voisinage de u_0 ($\mu_0(V_{u_0}) \neq 0$),

le dernier terme tend vers zéro quand n tend vers l'infini.

Le reste de la démonstration est alors analogue à la deuxième partie de la démonstration du théorème 3.2.

Soit $\hat{u}_{nh}$ l'estimateur bayésien approché à l'ordre. On peut alors énoncer :

Corollaire 4.1.1 : Sous les hypothèses du théorème 4.1, pour tout h et presque toute suite d'observation, l'estimateur $\hat{u}_{nh}$ converge vers la valeur vraie u_0 du paramètre à estimer quand n tend vers l'infini.

(pour la démonstration, voir celle du corollaire 3.2.1).

4.3

L'hypothèse " $Y(u)$ borné sur U_0 " n'est pas très restrictive. Nous allons le vérifier sur un des exemples étudiés numériquement.

Considérons le système gouverné par l'équation d'état suivante :

$$(I) \begin{cases} \frac{\partial y}{\partial t} - \frac{\partial}{\partial x} (g(x,u) \frac{\partial y}{\partial x}) = f & (x,t) \in Q =]0,1[\times]0,T[\\ y(0,t) = y(1,t) = 0 & \Omega =]0,1[\\ y(x,0) = y_0(x) \end{cases}$$

On suppose que f appartient à $L^2(Q)$, y_0 à $L^2(\Omega)$, et que g vérifie :

$$0 < \alpha \leq g(x,u) \leq M, \quad \forall x \in \Omega, \quad \forall u \in U_0$$

On sait que, dans ces conditions, la solution y de (I) est dans $W(0,T)$. En particulier $y(x,t;u)$ appartient à $L^2(0,T;H_0^1)$, pour tout u de U_0 ; c'est-à-dire : pour presque tout $t \in (0,T)$, $y(x,t;u) \in H_0^1(\Omega)$.

Or $H_0^1(\Omega)$ est contenu dans $C^0(\bar{\Omega})$ (muni de la norme du sup) avec injection continue (Cf. LIONS-MAGENES [8]). On a alors :

Soit $a \in \Omega$, pour presque tout t

$$|y(a,t;u)| \leq \|y(u)\|_{\infty} \leq C \|y(u)\|_{H_0^1} \quad \text{d'où :}$$

$$\|\bar{Y}(u)\|_K = \int_0^T |y(a,t;u)|^2 dt \leq C^2 \int_0^T \|y(u)\|_{H_0^1}^2 dt$$

Le terme de droite se majore à partir de l'inégalité de l'énergie.

L'équation d'état s'écrit, sous forme variationnelle :

$$\left(\frac{dy(t)}{dt}, y(t) \right)_{H^{-1}, H^1} + \int_{\Omega} g(x,u) \frac{\partial y}{\partial x} \frac{\partial y}{\partial x} dx = (f(t), y(t))_{H^{-1}, H^1}$$

d'où

$$\int_0^t \left(\frac{dy(t)}{dt}, y(t) \right)_{H^{-1}, H^1} dt + \alpha \int_0^t \left| \frac{\partial y}{\partial x} \right|_{L^2}^2 dt \leq \int_0^t (f(t), y(t))_{H^{-1}, H^1} dt$$

L'inégalité de Poincaré nous donne :

$$k \| y \|_{H_0^1(\Omega)} \leq \left| \frac{\partial y}{\partial x} \right|_{L^2(\Omega)}$$

d'où

$$\begin{aligned} \frac{1}{2} |y(t)|_{L^2}^2 + k^2 \alpha \int_0^T \|y(u)\|_{H_0^1}^2 &\leq \frac{1}{2} |y_0|_{L^2}^2 + \int_0^T \|f\|_{H^{-1}} \|y\|_{H_0^1} dt \\ &\leq \frac{1}{2} |y_0|_{L^2}^2 + K \int_0^T \|f\|_{H^{-1}}^2 dt + \frac{k^2 \alpha}{2} \int_0^T \|y\|_{H_0^1}^2 dt \end{aligned}$$

On obtient finalement :

$$\int_0^T \|y(u)\|_{H_0^1}^2 dt \leq \frac{2}{k^2 \alpha} \left(\frac{1}{2} |y_0|_{L^2}^2 + K \int_0^T \|f\|_{H^{-1}}^2 dt \right)$$

Le terme de droite est indépendant de u ; l'hypothèse H.3. est ainsi vérifiée.

III. APPLICATIONS NUMERIQUES

Les essais numériques ont porté sur l'identification des coefficients d'une équation aux dérivées partielles du type équation de la chaleur. Physiquement, le problème peut s'interpréter comme l'identification de la transmissivité thermique d'un fil conducteur, l'observation étant donnée par l'évolution de la température en un point du fil. Nous avons étudié les cas suivants :

- (i) le fil est homogène (du point de vue thermique) ; la transmissivité est constante le long du fil.
- (ii) la transmissivité est une fonction en escalier; le fil est constituée par plusieurs morceaux de transmissivités différentes.

(Pour le cas d'un fil de transmissivité " quelconque ", voir les applications numériques de la deuxième partie).

5. SYSTEMES ET APPROXIMATIONS.

5.1 Les systèmes

Nous avons étudié numériquement deux systèmes très voisins.

- (i) le système A, donné par les équations suivantes :

$$(I) \left\{ \begin{array}{l} \frac{\partial y}{\partial t} - u \frac{\partial^2 y}{\partial x^2} + v y = f \quad \text{sur } (0,1) \times (0,T) \\ \frac{\partial y}{\partial x}(0,t) = \frac{\partial y}{\partial x}(1,t) = 0 \\ y(x,0) = 0 \end{array} \right.$$

(u et v sont les paramètres à estimer).

L'observation est définie par :

$$(5.1) \quad z(t) = y(0,t;u_0, v_0) + \eta(t)$$

(u_0 et v_0 désignent la valeur vraie des paramètres u et v ,

$\eta(t)$ est caractérisé par une mesure cylindrique de Gauss sur $L^2(0,T)$ d'espérance mathématique nulle et d'opérateur de covariance $\sigma^2 I$)

(ii) Le système B.

$$\text{II} \quad \begin{cases} \frac{\partial y}{\partial t} - \frac{\partial}{\partial x} \left(g(x,u) \frac{\partial y}{\partial x} \right) = f & \text{sur } (0,1) \times (0,T) \\ y(0,t) = y(1,t) = 0 \\ y(x,0) = 0 \end{cases}$$

($g(x,u)$ est une fonction en escalier en x dépendant de $u \in \mathbb{R}^p$ qu'il s'agit d'identifier.)

L'observation est donnée par :

$$(5.2) \quad z(t) = y(b, t; u_0) + \eta(t)$$

(u_0 et η ont la même signification que précédemment, b est un nombre fixé de $]0,1[$ qui ne coïncide pas avec un point de discontinuité de g).

5.2 Approximation de l'espace des observations.

L'espace des observations est $L^2(0,T)$. On découpe l'intervalle $(0,T)$ en $N(h)$ parties égales ($h = \frac{T}{N(h)}$). Considérons le système ortho-normé de $L^2(0,T)$ donné par :

$$(5.3) \quad e_h^j = e_h^j(t) = \begin{cases} \frac{1}{\sqrt{h}} & \text{si } t \in [(j-1)h, jh] \\ 0 & \text{sinon} \end{cases} \quad j = 1, \dots, N(h)$$

L'espace K_h est alors le sous-espace de $K = L^2(0, T)$ engendré par les vecteurs e_h^j .

La projection de Y_d dans $\mathbb{R}^{N(h)}$ est définie par :

$$(5.4) \quad (\tau_h Y_d)_j = \frac{1}{\sqrt{h}} \int_{(j-1)h}^{jh} y(b, t; u_0) dt + \zeta_j$$

où ζ_j est la résolution d'une variable aléatoire normale centrée d'écart-type σ . On a de même :

$$(5.5) \quad (\tau_h \bar{Y}(u))_j = \frac{1}{\sqrt{h}} \int_{(j-1)h}^{jh} y(b, t; u) dt$$

En fait $\bar{Y}(u)$ n'est pas connu explicitement et sera approché par $\bar{Y}_\xi(u)$ à partir d'un schéma d'approximation des équations (I) ou (II).

Les différents estimateurs sont alors :

(i) estimateur bayésien approché :

$$(5.6) \quad \hat{u}_{h\xi}(\tau_h Y_d) = \frac{\int_{U_0} d\mu_0(u) u \exp \left[-\frac{1}{2\sigma^2} \sum_{j=1}^{N(h)} (\tau_h Y_d - \tau_h \bar{Y}_\xi(u))^2_j \right]}{\int_{U_0} d\mu_0(u) \exp \left[-\frac{1}{2\sigma^2} \sum_{j=1}^{N(h)} (\tau_h Y_d - \tau_h \bar{Y}_\xi(u))^2_j \right]}$$

(ii) estimateur bayésien récursif approché, avec observation unique:

$$(5.7) \quad \hat{u}_{nh\xi}(\tau_h Y_d) = \frac{\int_{U_0} d\mu_0(u) u \exp \left[-\frac{n}{2\sigma^2} \sum_{j=1}^{N(h)} (\tau_h Y_d - \tau_h \bar{Y}_\xi(u))^2_j \right]}{\int_{U_0} d\mu_0(u) \exp \left[-\frac{n}{2\sigma^2} \sum_{j=1}^{N(h)} (\tau_h Y_d - \tau_h \bar{Y}_\xi(u))^2_j \right]}$$

(iii) estimateur bayésien récursif approché, avec une suite d'observations :

$$(5.8) \quad \hat{u}_{nh}(X_1, X_2, \dots, X_n) = \frac{\int_{U_0} d\mu_0(u) u \exp \left[-\frac{n}{2\sigma^2} \sum_{j=1}^{N(h)} \left(\frac{X_1 + X_2 + \dots + X_n}{n} - \tau_h \bar{Y}_\xi(n) \right)_j^2 \right]}{\int_{U_0} d\mu_0(u) \exp \left[-\frac{n}{2\sigma^2} \sum_{j=1}^{N(h)} \left(\frac{X_1 + X_2 + \dots + X_n}{n} - \tau_h \bar{Y}_\xi(u) \right)_j^2 \right]}$$

où $X_1, X_2, \dots, X_n$ sont des vecteurs aléatoires (de dimension $N(h)$) gaussiens, de moyenne $\tau_h \bar{Y}(u_0)$ et de matrice de covariance $\sigma^2 I$.

6. RESULTATS NUMERIQUES.

La loi a priori μ_0 sera prise uniforme sur U_0 dans tous les cas.

6.1 Identification d'un paramètre.

On prend le système A avec $v = 0$ dans (I). Il reste donc à identifier le paramètre réel u . Choisissons pour second membre :

$$(6.1) \quad f(x, t) = \gamma \cos \pi x e^{\beta \pi^2 t}$$

(γ et β étant des constantes réelles)

La solution de l'équation (I) s'exprime alors analytiquement :

$$(6.2) \quad y(x, t; u) = \frac{\gamma \cos \pi x}{(u + \beta) \pi^2} (e^{\beta \pi^2 t} - e^{-u \pi^2 t})$$

Autres caractéristiques :

- valeur vraie du paramètre : $u_0 = 7$
- support de la loi a priori : $U_0 = (1, 51)$
- discrétisation de U_0 ; pas en $u = 1$.

- discrétisation de l'équation aux dérivées partielles (I) :

$$\text{pas en temps} = 1/20$$

$$\text{pas en espace} = 1/40$$

- dimension de K_h : $N(h) = 20$, $T = 1$ ($h = 1/20$)

- on fixe $\beta = 0.5$

On fera des essais pour différentes valeurs de γ et σ .

Résultats.

6.1.1 Sans récursivité.

On calcule l'estimateur (5.6) que l'on désigne ici par $\hat{u}$.

1. Avec $\begin{cases} \gamma = 10 \\ \sigma = 0.1 \end{cases}$ on obtient $\hat{u} = 7$

Toute la masse de probabilité (de la loi a posteriori) est concentrée au point $u = 7$.

2. Avec $\begin{cases} \gamma = 1 \\ \sigma = 0.1 \end{cases}$ on obtient $\hat{u} = 6.73$

3. Avec $\begin{cases} \gamma = 1 \\ \sigma = 1 \end{cases}$ on a $\hat{u} = 23.12$

La densité a posteriori des cas 2 et 3 est représentée sur les figures 1 et 2 respectivement.

Commentaire.

Sur ces exemples, la qualité de l'estimation dépend fortement des paramètres γ et σ . Il est normal que cette qualité diminue lorsque σ augmente, ce qui signifie que le bruit devient de plus en plus important. La dépendance vis à vis de γ est naturelle également; l'équation (I) est linéaire, ainsi si γ (c'est-à-dire le second membre) augmente la solution y croît , la valeur relative du bruit diminue donc, d'où une meilleure estimation. Pour identifier correctement le rapport entre σ et γ doit être de 1 à 10 (au moins) ce qui correspond à une valeur relative du bruit de l'ordre de 5 % .

Dans le cas le plus défavorable (3), si l'estimateur $\hat{u}$ nous donne peu de renseignements ($\hat{u}$ est voisin du milieu de U_0) la densité a posteriori (figure 2) indique tout de même que la valeur du paramètre à estimer u est probablement comprise entre 1 et 10. La récursivité sera intéressante dans ce cas, nous le verrons dans les essais suivants.

6.1.2 Récursivité avec observation unique.

On utilise ici la formule (5.7) et on désigne l'estimateur récursif par $\hat{u}_n$. Le problème consiste à choisir n , c'est-à-dire déterminer l'ordre pour lequel la loi a posteriori sera suffisamment "pointue". Nous avons procédé de la façon suivante : si l'écart-type e_1 de la loi a posteriori (à l'ordre 1) est plus petit que 0.5, on prend $n = 1$, sinon on fait $n = 10$ et :

- si $e_n \leq 0.5$ on conserve n et on s'arrête
- sinon on multiplie n par 5 et on recommence.

Nous avons obtenu les résultats suivants :

1. Avec $\gamma = 1$, $\sigma = 0.1$.

Série de 10 essais.

Essai n°	1	2	3	4	5	6	7	8	9	10
n	10	10	50	10	10	10	10	10	50	50
Estimateur $\hat{u}_n$	6.08	7.19	8.01	5.85	6.85	6.01	5.95	6.94	7.60	9.
Ecart-type e_n	0.27	0.41	0.09	0.36	0.38	0.10	0.22	0.31	0.49	0

Moyenne empirique : 7.55

Ecart-type empirique : 2.77

Sur 50 essais, on obtient :

Moyenne empirique : 7.13

Ecart-type empirique : 1.33

2. Avec $\gamma = 1$, $\sigma = 0.5$.

Série de 10 essais.

Essai N°	1	2	3	4	5	6	7	8	9	10
n	50	1250	31250	10	250	50	10	250	1250	50
Estimateur $\hat{u}_n$	4.14	8.01	23.10	3.15	5.93	3.98	3.25	6.01	10.34	3.99
Ecart-type e_n	0.35	0.07	0.31	0.40	0.26	0.15	0.49	0.12	0.47	0.12

Moyenne empirique : 7.19

Ecart-type empirique : 5.72

Sur 50 essais :

Moyenne empirique : 6.43

Ecart-type empirique : 3.7

3. Avec $\gamma = 1$, $\sigma = 1$.

Série de 10 essais.

Essai N°	1	2	3	4	5	6	7	8	9	10
n	50	6250	6250	10	250	250	10	250	156250	1250
Estimateur $\hat{u}_n$	3.08	9.95	5.10	2.10	4.89	2.40	2.17	6.02	19.07	50.81
Ecart-type e_n	0.29	0.22	0.20	0.37	0.36	0.49	0.46	0.49	0.25	0.47

Moyenne empirique : 10.5

Ecart-type empirique : 14.32

Si on élimine l'essai N° 10 qui fournit un estimateur à la limite de U_0 ,
on a :

Moyenne empirique : 6.09

Ecart-type empirique : 5.16

Commentaire

La récursivité avec observation unique est, d'une part, très simple à mettre en oeuvre. D'autre part, elle permet d'obtenir des renseignements lorsque l'estimateur ordinaire n'en donne pas. C'est le cas de la troisième série d'essais. Alors que l'estimateur à l'ordre 1 dépend essentiellement de U_0 (Il est voisin du milieu de l'intervalle U_0), à partir de l'estimateur récursif, on peut conclure que la valeur vraie du paramètre u avoisine 6 et qu'elle est certainement comprise entre 1 et 11.

6.2

Identification de deux paramètres.

On prend encore le système A , avec pour second membre :

$$f(x,t) = \gamma e^{xt}$$

Dans ce cas, nous n'avons pas d'expression analytique de la solution; pour générer l'observation, on utilisera une résolution numérique très fine de l'équation (I).

Autres caractéristiques :

- valeurs vraies des paramètres : $u_0 = 15$, $v_0 = 7$
- support de la loi a priori : $U_0 = (1,20) \times (1,20)$
- discrétisation de U_0 : pas en u = pas en v = 1 .

Les autres caractéristiques sont analogues à celles de l'exemple précédent.

Résultat.

Avec $\gamma = 10$ et $\sigma = 0.01$, on obtient :

$$\hat{u} = 14.82 \quad \hat{v} = 7$$

La masse de probabilité (de la loi a posteriori) est concentrée le long de la droite $v = 7$; on a représenté la coupe de la densité

pour $v = 7$ sur la figure 3.

Nous avons fait une autre série d'essais avec des caractéristiques légèrement différentes :

- valeur vraie des paramètres : $u_0 = 25$ $v_0 = 14$

- support de μ_0 : $U_0 = (1,40) \times (1,40)$

- discrétisation : pas en u = pas en v = 1 .

Résultats :

Pour un couple de valeurs (γ, σ) on fait 50 essais ; le tableau ci-dessous donne les estimateurs moyens (ou moyenne empirique des estimateurs).

γ	σ	$\hat{u}$ moyen	$\hat{v}$ moyen	
1	0.01	20.94	14.87	
5	0.01	22.18	13.96	
10	0.01	24.36	13.97	(*)
1	0.05	20.54	22.41	
5	0.05	20.91	15.02	
10	0.05	20.75	14.10	(**)
1	0.1	20.49	22.00	
5	0.1	20.87	18.01	
10	0.1	20.88	14.95	

Sur la figure 4 , on a tracé les ellipses de concentration des cas (*) et (**)

Commentaire.

Il ressort de ces essais que, dans l'identification de plusieurs

paramètres (ou d'un paramètre vectoriel), la qualité de l'estimateur peut différer fortement selon le paramètre (ou la composante) estimé. Dans l'exemple traité ci-dessus, on estime beaucoup mieux v que u . Donnons quelques explications.

Avec les résultats de CHAVENT [4] on peut établir les formules suivantes :

$$\frac{\partial \Psi}{\partial u} = \int_Q \frac{\partial y}{\partial x} \frac{\partial p}{\partial x} dx dt \quad (Q = (0,1) \times (0,T))$$

$$\frac{\partial \Psi}{\partial v} = \int_Q y p dx dt$$

p étant solution de :

$$\begin{cases} -\frac{\partial p}{\partial t} - u \frac{\partial^2 p}{\partial x^2} + vp = -2\delta(x) (y(0,t;u,v) - Y_d(t)) \\ \frac{\partial p}{\partial x}(0,t) = \frac{\partial p}{\partial x}(1,t) = 0 \\ p(T) = 0 \end{cases}$$

Quand y est solution d'un problème de Neumann (c'est le cas du système A), il est fréquent que y varie peu avec x . Ainsi y est nettement plus grand que sa dérivée $\frac{\partial y}{\partial x}$, d'où $\frac{\partial \Psi}{\partial v}$ est nettement supérieur à $\frac{\partial \Psi}{\partial u}$. Cela explique que la densité a posteriori dépende plus fortement de v que de u . Notons que dans le cas d'un problème de Dirichlet, on a le phénomène inverse, Ψ dépend plus fortement de u que de v .

6.3 Identification de deux paramètres. Autre exemple.

On considère un système du type B cette fois. Le coefficient $g(x,u)$ sera une fonction en escalier à deux marches :

$$g(x,u) = \begin{cases} u_1 & \text{sur } (0, 1/2) \\ u_2 & \text{sur } (1/2, 1) \end{cases}$$

On choisit le second membre f tel que la solution de (II) s'exprime analytiquement ; avec :

$$f(x,t) = \begin{cases} (u_{10} + 2u_{20}t)(2(\frac{1}{u_{10}} - 1)x^2 + (2 - \frac{1}{u_{10}})x) + (4u_{10} - 4)(u_{10}t + u_{20}t^2) & \text{sur } (0, \frac{1}{2}) \\ (u_{10} + 2u_{20}t)(-2(1 + \frac{1}{u_{20}})x^2 + (2 + \frac{3}{u_{20}})x - \frac{1}{u_{20}}) + (4u_{20} + 4)(u_{10}t + u_{20}t^2) & \text{sur } (\frac{1}{2}, 1) \end{cases}$$

on obtient :

$$y(x,t) = \begin{cases} (u_{10}t + u_{20}t^2)(2(\frac{1}{u_{10}} - 1)x^2 + (2 - \frac{1}{u_{10}})x) & \text{sur } (0, \frac{1}{2}) \\ (u_{10}t + u_{20}t^2)(-2(1 + \frac{1}{u_{20}})x^2 + (2 + \frac{3}{u_{20}})x - \frac{1}{u_{20}}) & \text{sur } (\frac{1}{2}, 1) \end{cases}$$

(u_{10} et u_{20} sont les valeurs vraies des paramètres à estimer).

Autres caractéristiques :

- on prend $b = \frac{1}{4}$ dans (5.2) .
- support de la loi a priori : $U_0 = (1,20) \times (1,20)$
- discrétisation de U_0 : pas en u_1 = pas en u_2 = 1
- discrétisation de l'équation aux dérivées partielles (II) :
pas en temps = pas en espace = $\frac{1}{40}$
- dimension de K_h : $N(h) = 40$ ($T = 1$, $h = \frac{1}{40}$)

Résultats.

6.3.1 Sans récursivité.

On désignera par $\hat{u}_1$ et $\hat{u}_2$ les estimateurs de u_1 et u_2 , σ_1 et σ_2 la variance (respectivement en u_1 et u_2) de la loi a posteriori.

6.3.1.1. Avec $u_{10} = 7$, $u_{20} = 11$.

$\sigma = 0.01$	$\hat{a}_1 = 7.14$ $\hat{a}_2 = 11.23$	$\sigma_1 = 1.4$ $\sigma_2 = 3.3$
$\sigma = 0.05$	$\hat{a}_1 = 8.28$ $\hat{a}_2 = 9.57$	$\sigma_1 = 2.7$ $\sigma_2 = 5.0$
$\sigma = 0.1$	$\hat{a}_1 = 8.52$ $\hat{a}_2 = 9.31$	$\sigma_1 = 2.8$ $\sigma_2 = 5.2$

6.3.1.2. Avec $u_{10} = 3$, $u_{20} = 17$.

$\sigma = 0.01$	$\hat{a}_1 = 5.06$ $\hat{a}_2 = 13.61$	$\sigma_1 = 3.0$ $\sigma_2 = 4.7$
$\sigma = 0.05$	$\hat{a}_1 = 8.37$ $\hat{a}_2 = 8.74$	$\sigma_1 = 3.6$ $\sigma_2 = 5.0$

6.3.2. Récursivité avec une suite d'observations.

On utilise ici la formule (5.8)

6.3.2.1. Avec $u_{10} = 7$, $u_{20} = 11$.

$\sigma = 0.01$									
n	1	2	3	5	10	20	30	50	100
$\hat{u}_{1n}$	7.14	7.17	7.15	7.10	7.03	7.002	7	7	7
$\hat{u}_{2n}$	11.23	10.82	10.78	10.81	10.95	10.997	11	11	11
σ_{1n}	1.4	0.8	0.6	0.3	0.2	0.04	0.01	0.001	0
σ_{2n}	3.3	2.0	1.5	0.7	0.3	0.08	0.02	0.002	0

$\sigma = 0.05$									
n	1	2	3	5	10	20	30	50	100
$\hat{u}_{1n}$	8.28	7.96	8.13	8.06	8.39	7.65	7.67	7.58	7.47
$\hat{u}_{2n}$	9.57	9.99	9.74	9.86	9.18	10.22	9.99	9.93	10.07
σ_{1n}	2.7	2.5	2.5	2.4	2.4	1.7	1.5	1.0	0.5
σ_{2n}	5.0	4.7	4.8	4.7	4.6	3.5	2.9	1.8	1.0

$\sigma = 0.1$									
n	1	5	10	50	100	200	300	400	500
$\hat{u}_{1n}$	8.52	8.37	8.87	8.62	8.22	7.72	7.31	7.41	7.30
$\hat{u}_{2n}$	9.31	9.41	8.50	8.52	8.96	9.63	10.40	10.18	10.42
σ_{1n}	2.8	2.7	2.8	2.2	1.8	0.9	0.5	0.5	0.4
σ_{2n}	5.2	5.0	5.0	4.1	3.0	1.5	1.2	1.1	0.9

6.3.2.2. Avec $u_{10} = 3$, $u_{20} = 17$

n	$\sigma = 0.01$				$\sigma = 0.05$			
	$\hat{u}_{1n}$	$\hat{u}_{2n}$	σ_{1n}	σ_{2n}	$\hat{u}_{1n}$	$\hat{u}_{2n}$	σ_{1n}	σ_{2n}
1	5.06	13.61	3.0	4.7	8.37	8.74	3.6	5.0
5	3.09	16.84	0.7	1.2	8.30	8.72	3.5	5.0
10	3.03	16.95	0.4	0.7	7.20	10.20	3.4	5.0
50	3	17	10^{-5}	10^{-4}	4.50	14.52	2.7	4.3
100	3	17	0	0	3.47	16.10	1.3	2.3

Commentaire.

Sur ces exemples, il apparaît encore que la loi a posteriori (à l'ordre 1) diffère peu de la loi a priori lorsque le bruit prend des proportions trop importantes. La récursivité (avec une suite d'observations) permet alors d'obtenir des résultats ; on est toutefois limité par le nombre d'observations nécessaires qui croît rapidement avec l'écart-type du bruit. De façon plus précise, pour obtenir des résultats valables, la valeur relative du bruit doit être :

- (i) inférieure à $1/1000$ pour l'estimateur ordinaire.
- (ii) inférieure à $1/100$ pour l'estimateur récursif avec un nombre raisonnable d'observations (de 10 à 50).

Notons un fait désagréable : la convergence des estimateurs récursifs n'est pas monotone, ainsi quelques observations supplémentaires n'apportent pas nécessairement une amélioration de l'estimation. Il nous reste à parler des temps de calcul. D'une manière générale, la méthode bayésienne demande des temps assez longs ; ils ne sont cependant pas excessifs dans les exemples précédents (moins de 5mn d'UNIVAC 1108).

6.4 Identification de quatre paramètres.

On prend encore le système B , mais avec une fonction en escalier $g(x,u)$ à quatre marches :

$$g(x,u) = \begin{cases} u_1 & \text{sur } (0, 1/4) \\ u_2 & \text{sur } (1/4, 1/2) \\ u_3 & \text{sur } (1/2, 3/4) \\ u_4 & \text{sur } (3/4, 1) \end{cases}$$

et, comme second membre :

$$f(x,t) = -2x(x-1)(5+6t) + 8(5t+3t^2)$$

Pour simuler l'observation, on utilise une approximation très fine de l'équation (II).

Autres caractéristiques :

- abscisse de l'observation : $b = 3/8$
- valeur vraie des paramètres : $u_{10} = 0.06$
 $u_{20} = 0.08$
 $u_{30} = 0.12$
 $u_{40} = 0.14$
- support de la loi a priori : $U_0 = (0.02, 0.2)^4$
- discrétisation de U_0 : pas en $u_i = 0.02$ $i = 1, 2, 3, 4$.
- discrétisation de l'équation (II) :

pas en temps = pas en espace = $1/80$
- dimension de K_h : $N(h) = 80$ ($T = 1$, $h = 1/80$)

Résultats : récursivité avec une suite d'observations.

6.4.1. Avec $\sigma = 0.01$

n	$\hat{u}_{1n}$	$\hat{u}_{2n}$	$\hat{u}_{3n}$	$\hat{u}_{4n}$
1	0.071	0.098	0.108	0.113
5	0.085	0.060	0.094	0.115
10	0.073	0.070	0.095	0.134
20	0.065	0.075	0.113	0.134
30	0.061	0.079	0.118	0.140
50	0.0602	0.0806	0.118	0.1415

6.4.2. Avec $\sigma = 0.05$

n	$\hat{u}_{1n}$	$\hat{u}_{2n}$	$\hat{u}_{3n}$	$\hat{u}_{4n}$
1	0.086	0.088	0.102	0.090
10	0.145	0.032	0.116	0.117
50	0.120	0.041	0.129	0.117
100	0.092	0.050	0.135	0.138
150	0.071	0.057	0.135	0.156
200	0.066	0.061	0.124	0.155

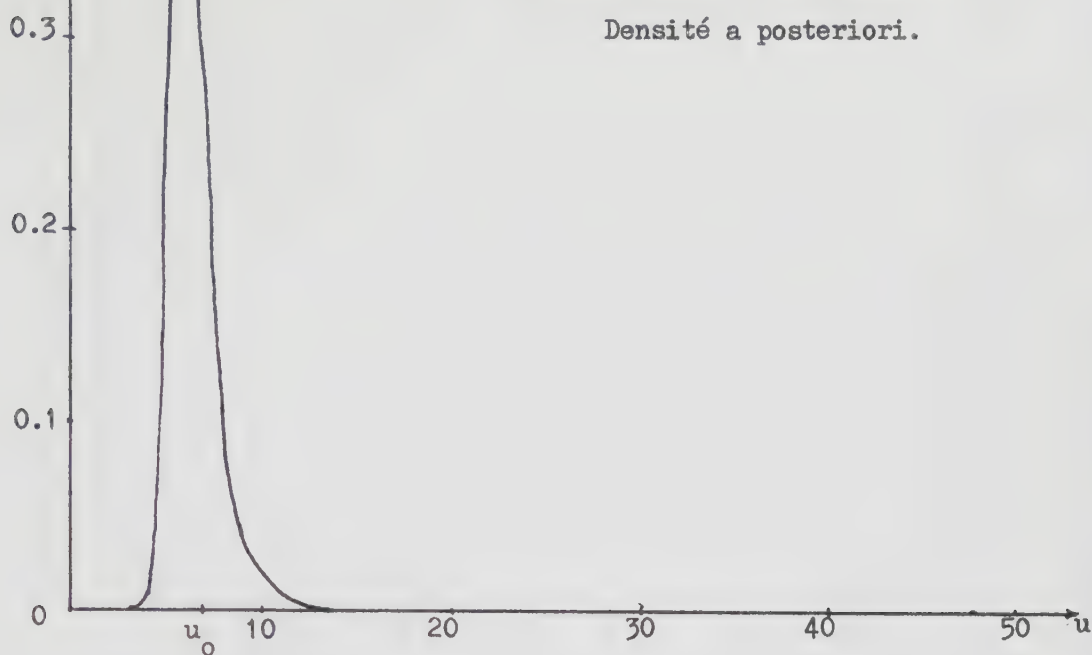
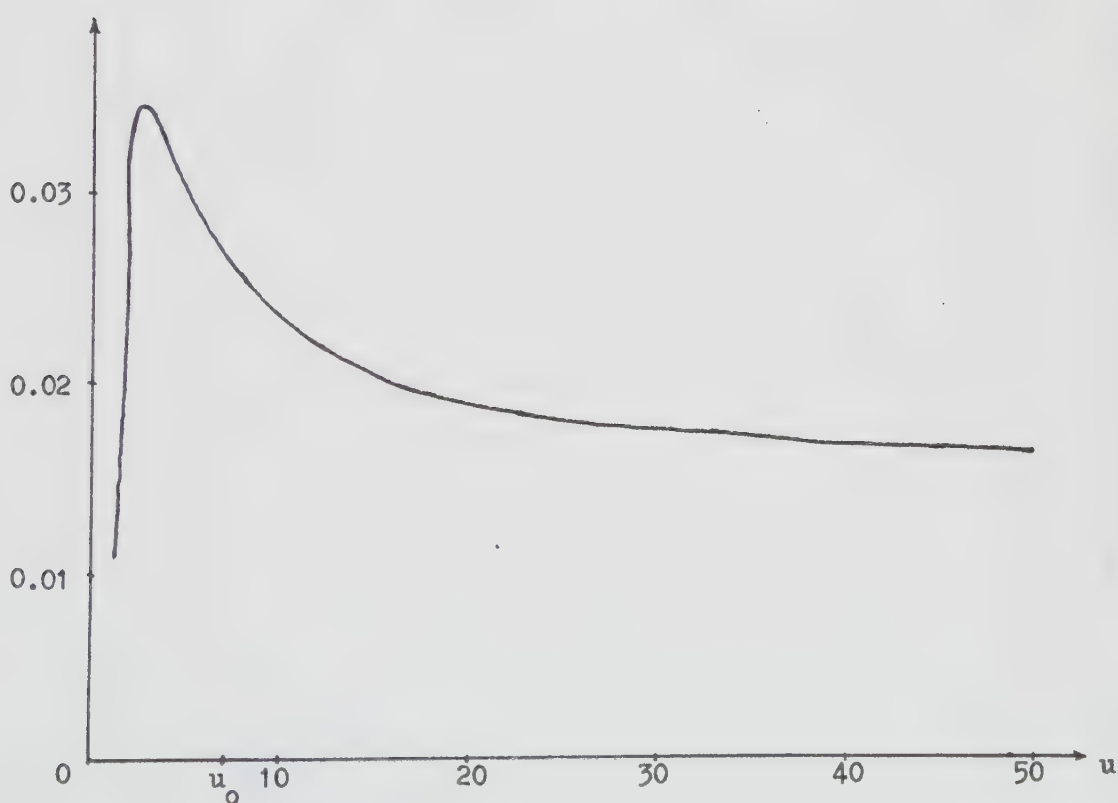
Commentaire.

On peut faire les mêmes remarques que pour l'exemple précédent, mis à part le temps de calcul. En effet, d'une part, on identifie 4 paramètres sur $U_0 = (1,10)^4$ qui est approché par un maillage à $9^4 (= 6561)$ points. Il faut donc résoudre l'équation (II) 6561

fois. D'autre part, la solution y varie peu lorsqu'on fait varier un seul paramètre d'une unité dans $g(x,u)$. Il faut alors une discrétisation assez fine pour résoudre l'équation (II) (80×80 dans notre exemple), d'où encore une augmentation du temps de calcul. On atteint alors 30 mn d'UNIVAC 1108, ce qui interdit d'utiliser cette méthode pour identifier plus de 2 paramètres. Dans la seconde partie, on présente d'autres techniques qui permettent d'identifier des paramètres distribués (ou à constante non localisée).

Identification d'un paramètre.

Densité a posteriori.

Figure 1 : $\sigma = 0.1$ Figure 2 : $\sigma = 1$

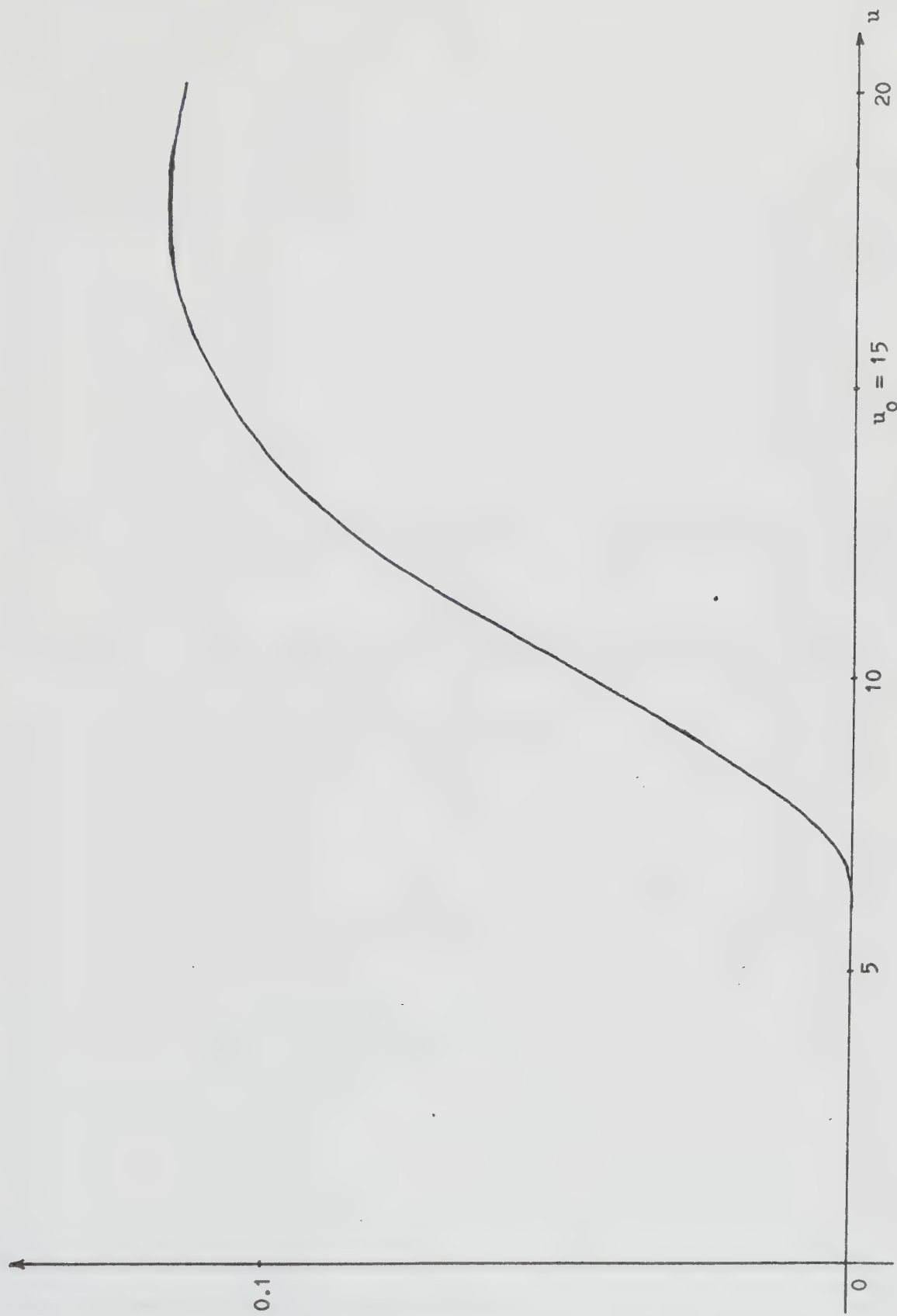


Figure 3 : Identification de deux paramètres. Coupe de la densité a posteriori pour $u = v_0 = 7$ avec $= 001$

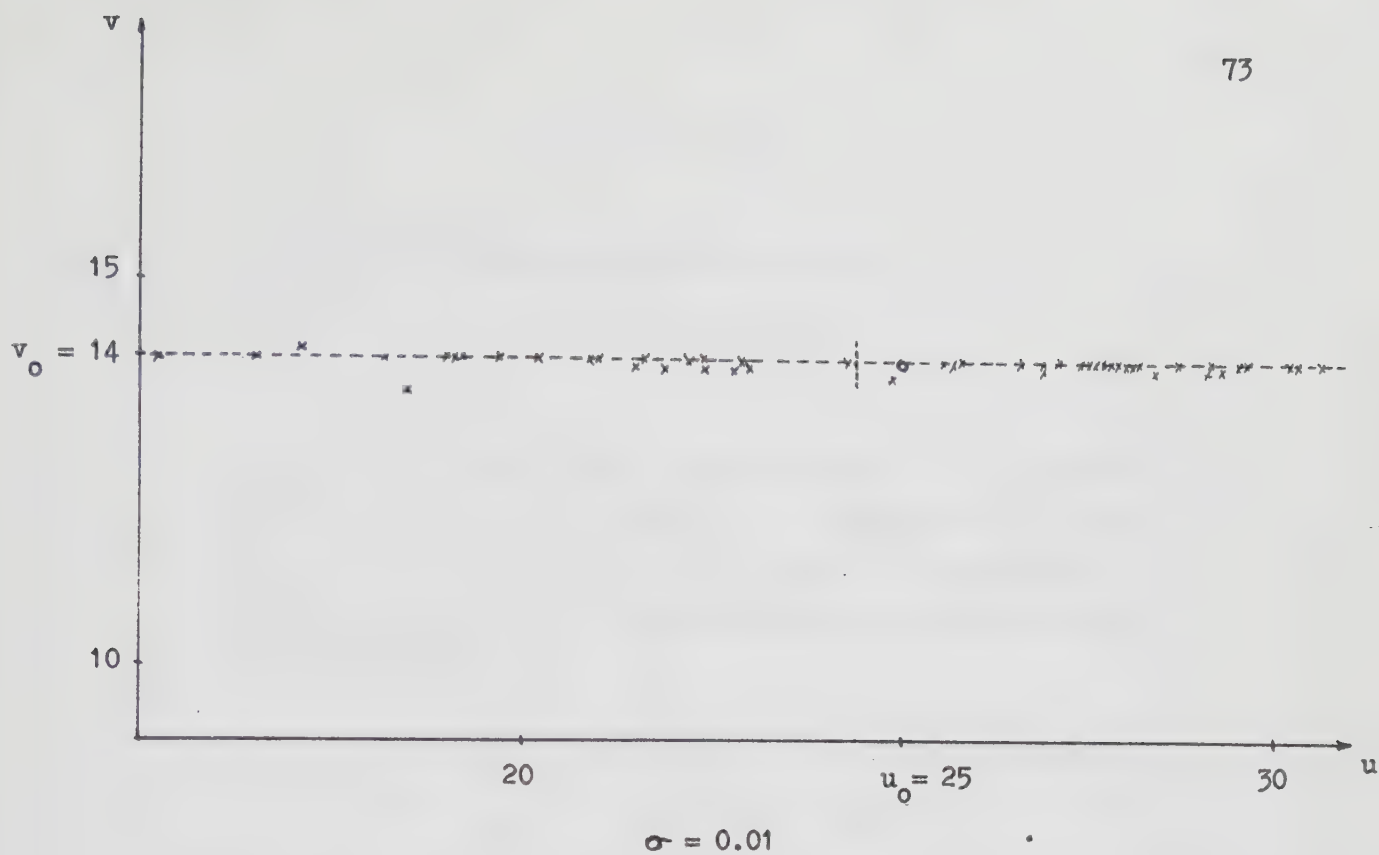
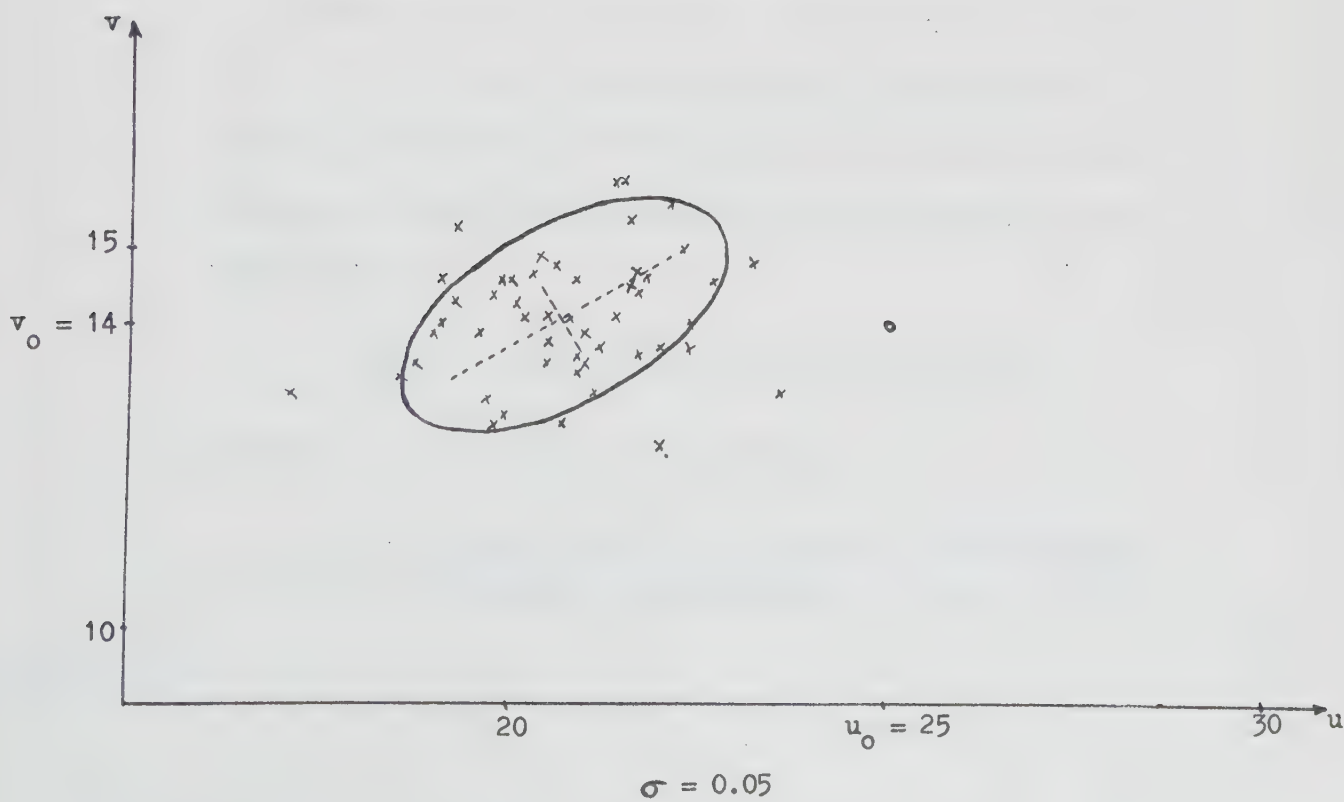


Figure 4 : Identification de deux paramètres. Ellipse de concentration (pour 50 essais)



DEUXIEME PARTIE

--

LA METHODE DES MOINDRES CARRES

UN ALGORITHME RECURSIF

Nous allons étudier dans cette partie un algorithme proposé pour les systèmes discrets (cet algorithme peut se généraliser aux systèmes différentiels, mais non aux systèmes aux dérivées partielles, nous l'appliquerons donc directement aux systèmes discrétisés).

Considérons un tel système, régi par les équations :

$$y_{n+1} = f(y_n, n) + (\text{bruit})$$

$$z_n = H_n y_n + (\text{bruit})$$

Le problème consiste à retrouver l'état y_n à partir de l'observation z_n , la fonction f étant non linéaire. L'estimé des moindres carrés est défini comme la trajectoire qui minimise une fonction quadratique des bruits. A partir des multiplicateurs de LAGRANGE et par linéarisation, on obtient l'algorithme suivant :

$$\hat{y}_{n+1} = f(\hat{y}_n, n) + C_n (z_{n+1} - H_{n+1} f(\hat{y}_n, n))$$

la matrice C_n étant définie par récurrence.

Il est aisé ensuite de se ramener à l'identification par addition de composantes supplémentaires au vecteur d'état.

1. ESTIMATION AU SENS DES MOINDRES CARRES

1.1

Position du problème.

Considérons un système discret régi par l'équation :

$$(1.1) \quad y_{n+1} = f(y_n, n) + (\text{bruit dynamique})$$

l'observation étant donnée par :

$$(1.2) \quad z_n = H_n y_n + (\text{erreur de mesure})$$

(y_n est un vecteur de dimension m , f une fonction vectorielle de dimension m , z_n un vecteur de dimension q , H_n une matrice $q \times m$). On désignera le bruit dynamique par $r_n^{(1)}$, l'erreur de mesure par $r_n^{(2)}$. Les équations (1.1) et (1.2) s'écrivent alors:

$$(1.3) \quad y_{n+1} = f(y_n, n) + r_n^{(1)}$$

$$(1.4) \quad z_n = H_n y_n + r_n^{(2)}$$

Etant donnée une suite d'observations z_n , $n = 0, 1, \dots, N$, on appelle trajectoire (ou estimée) des moindres carrés la suite $\hat{y}_n$, $n = 0, 1, \dots, N+1$, solution de (1.3) et (1.4), qui minimise la quantité

$$\sum_{n=0}^N \left\{ \frac{1}{2} \|r_n^{(1)}\|_{R_n}^2 + \frac{1}{2} \|r_n^{(2)}\|_{\varphi_n}^2 \right\} . \text{ Les matrices } R_n \text{ et } \varphi_n$$

sont des matrices symétriques positives $m \times m$ et $q \times q$ qui servent à définir la norme des bruits :

$$(1.5) \quad \|x\|_{R_n}^2 = x R_n x$$

$$(1.6) \quad \|x\|_{\varphi_n}^2 = x \varphi_n x$$

On ramène ainsi l'estimation à un problème d'optimisation :

trouver $y_n, r_n^{(1)}, n = 0, 1, \dots, N$ ($N+1$ pour y_n) qui minimise

$$(1.7) \quad I(y, r^{(1)}) = \sum_{n=0}^N \left\{ \frac{1}{2} \|r_n^{(1)}\|_{R_n}^2 + \frac{1}{2} \|z_n - H_n y_n\|_{\varphi_n}^2 \right\}$$

sous les contraintes :

$$(1.8) \quad y_{n+1} = f(y_n, n) + r_n^{(1)} \quad n = 0, 1, \dots, N$$

Ce problème conduit immédiatement à prendre

$$(1.9) \quad r_N^{(1)} = 0$$

1.2 Conditions nécessaires d'optimalité.

Elles découlent de la théorie des multiplicateurs de LAGRANGE. Si $y_n, n = 0, \dots, N+1, r_n^{(1)}, n = 0, \dots, N$ réalisent le minimum de $I(y, r^{(1)})$, il existe $(N+2)$ constantes $\lambda_{-1}, \lambda_0, \dots, \lambda_N$ telles que :

$$(\text{si } g_n(y_{n+1}, y_n, r_n^{(1)}) = f(y_n, n) + r_n^{(1)} - y_{n+1})$$

$$(1.10) \quad \left\{ \begin{array}{l} \frac{\partial I}{\partial y_n} - \sum_{j=-1}^N \lambda_j \frac{\partial g_j}{\partial y_n} = 0 \quad n = 0, 1, \dots, N+1 \\ \frac{\partial I}{\partial r_n^{(1)}} - \sum_{j=-1}^N \lambda_j \frac{\partial g_j}{\partial r_n^{(1)}} = 0 \quad n = 0, 1, \dots, N \end{array} \right.$$

Ce qui donne :

$$(1.12) \quad \lambda_{n-1} = \frac{\partial f}{\partial y_n} \lambda_n + H_n' \varphi_n (z_n - H_n y_n)$$

$$n = 0, 1, \dots, N$$

$$(1.13) \quad \lambda_n = R_n r_n^{(1)}$$

$$n = 0, 1, \dots, N$$

(1.9) et (1.13) entraîne :

$$(1.14) \quad \lambda_N = 0$$

En utilisant les contraintes (1.8), on obtient finalement le système :

$$(1.15) \quad \left\{ \begin{array}{l} y_{n+1} = f(y_n, n) + R_n^{-1} \lambda_n \\ \lambda_{n-1} = \frac{\partial f'}{\partial y_n} \lambda_n + H_n' (\varphi_n(z_n - H_n y_n)) \end{array} \right\} \quad n = 0, 1, \dots, N$$

$$(1.12) \quad (I) \quad \left\{ \begin{array}{l} \lambda_{n-1} = \frac{\partial f'}{\partial y_n} \lambda_n + H_n' (\varphi_n(z_n - H_n y_n)) \\ \lambda_N = 0 \end{array} \right.$$

$$(1.14) \quad \left\{ \begin{array}{l} \lambda_N = 0 \end{array} \right.$$

1.3

Résolution du système (I)

Nous allons considérer le système plus général suivant :

$$(1.16) \quad \left\{ \begin{array}{l} y_{n+1} = g(y_n, \lambda_n, n) \\ \lambda_{n-1} = h(y_n, \lambda_n, n) \end{array} \right\} \quad n = 0, 1, \dots, M$$

$$(1.17) \quad \left\{ \begin{array}{l} \lambda_{n-1} = h(y_n, \lambda_n, n) \\ \lambda_M = b \end{array} \right.$$

$$(1.18) \quad \left\{ \begin{array}{l} \lambda_M = b \end{array} \right.$$

On le désignera par système (M, b) ou $S(M, b)$.

Soient $[y_n^{(1)}(b, N), \lambda_{n-1}^{(1)}(b, N), n = 0, 1, \dots, N+1]$ la solution du système (N, b) et $[y_n^{(2)}(b, N+1), \lambda_{n-1}^{(2)}(b, N+1), n = 0, 1, \dots, N+2]$ celle du système $(N+1, b)$.

$$(1.19) \quad \text{Posons} \quad Z_N(b) = y_N^{(1)}(b, N)$$

$$(1.20) \quad Z_{N+1}(b) = y_{N+1}^{(2)}(b, N+1)$$

Nous nous proposons de trouver une relation entre $Z_{N+1}(b)$ et $Z_N(b)$, c'est-à-dire déterminer la modification qu'apporte une observation supplémentaire sur l'estimée à l'instant N .

$$(1.21) \quad \text{Posons} \quad y_n^{(2)} = y_n^{(1)} + \alpha_n \quad n = 0, 1, \dots, N+1$$

$$(1.22) \quad \lambda_n^{(2)} = \lambda_n^{(1)} + \beta_n \quad n = 1, 0, \dots, N$$

Alors :

$$\begin{aligned} \alpha_{n+1} &= y_{n+1}^{(2)} - y_{n+1}^{(1)} \\ &= g(y_n^{(2)}, \lambda_n^{(2)}, n) - g(y_n^{(1)}, \lambda_n^{(1)}, n) \end{aligned}$$

Par linéarisation, on obtient :

$$(1.23) \quad \alpha_{n+1} = \frac{\partial g}{\partial y_n}(y_n^{(1)}, \lambda_n^{(1)}) \alpha_n + \frac{\partial g}{\partial \lambda_n}(y_n^{(1)}, \lambda_n^{(1)}) \beta_n$$

De même pour β_n :

$$(1.24) \quad \beta_{n-1} = \frac{\partial h}{\partial y_n}(y_n^{(1)}, \lambda_n^{(1)}) \alpha_n + \frac{\partial h}{\partial \lambda_n}(y_n^{(1)}, \lambda_n^{(1)}) \beta_n$$

$$(1.25) \quad \text{Posons enfin} \quad \alpha_n = P_n \beta_{n-1}$$

Evaluons P_{n+1} en fonction de P_n .

$$\alpha_{n+1} = P_{n+1} \beta_n = \frac{\partial g}{\partial y_n} \alpha_n + \frac{\partial g}{\partial \lambda_n} \beta_n$$

d'après (1.25) et (1.24) :

$$\alpha_n = P_n \beta_{n-1} = P_n \frac{\partial h}{\partial y_n} \alpha_n + P_n \frac{\partial h}{\partial \lambda_n} \beta_n$$

Alors :

$$(1.26) \quad P_{n+1} = \frac{\partial g}{\partial y_n} (I - P_n \frac{\partial h}{\partial y_n})^{-1} P_n \frac{\partial h}{\partial \lambda_n} + \frac{\partial g}{\partial \lambda_n}$$

On peut alors exprimer $Z_{N+1}^{(b)}$ en fonction de $Z_N^{(b)}$.

(1.20), (1.21), (1.16) et (1.19) entraîne

$$\begin{aligned}
 Z_{N+1}(b) &= g(Z_N(b), b, N) + \alpha_{N+1} \\
 &= g(Z_N(n), b, N) + P_{N+1} \beta_N
 \end{aligned}$$

(1.22) et (1.17) entraîne

$$\beta_N = h(Z_{N+1}(b), b, N+1) - b$$

d'où

$$(1.27) \quad Z_{N+1}(b) = g(Z_N(b), b, N) + P_{N+1} (h(Z_{N+1}(b), b, N+1) - b)$$

1.4

Algorithme récursif.

Revenons à notre problème initial pour évaluer les différents termes de (1.27). Supposons connue la trajectoire des moindres carrés à l'ordre N , soit $\hat{y}_N$ sa valeur à l'instant N . A partir de (1.27), on détermine la valeur à l'instant $(N+1)$, de l'estimée à l'ordre $(N+1)$. (Par estimée à l'ordre N , on entend estimée à partir des observations $z_0, z_1, \dots, z_N$).

Dans (1.27), on prendra :

$$b = 0$$

$$g(Z_N(b), b, N) = f(\hat{y}_N, N)$$

$$h(Z_{N+1}(b), b, N+1) = H_{N+1}' \varphi_{N+1}(z_{N+1} - H_{N+1} \hat{y}_{N+1})$$

On fera encore une approximation du premier ordre en remplaçant $\hat{y}_{N+1}$ par $f(\hat{y}_N, N)$. On obtient alors :

$$(1.28) \quad \hat{y}_{N+1} = f(\hat{y}_N, N) + P_{N+1} H_{N+1}' \varphi_{N+1}(z_{N+1} - H_{N+1} f(\hat{y}_N, N))$$

avec

$$(1.29) \quad P_{N+1} = \frac{\partial f}{\partial y_N}(\hat{y}_N) \left[I + P_N H_N' \varphi_N H_N \right]^{-1} P_N \frac{\partial f'}{\partial y_N}(\hat{y}_N) + R_N^{-1}$$

Ces formules doivent être complétées par la valeur initiale de $\hat{y}_N$ et $P_N \cdot \hat{y}_0$ est l'estimée de l'état à l'origine. Cette estimation pourra être assez grossière sans être trop éloignée de la solution exacte cependant, sinon le procédé récursif diverge. Le choix de P_0 est plus arbitraire. Dans le second membre de (1.28), le premier terme peut être considéré comme une prédiction, le second comme une correction à partir de l'observation; P_N s'interprète alors comme le gain affectant le terme correcteur. La valeur de P_0 reflètera donc la confiance que l'on accorde à l'estimation initiale $\hat{y}_0$, c'est-à-dire qu'il faut choisir un couple $(\hat{y}_0, P_0)$ et non pas chaque élément séparément. En effet, si $\hat{y}_0$ est proche de la valeur exacte, une valeur trop importante de P_0 risque d'éloigner $\hat{y}_1$ au point de faire diverger l'algorithme. Inversement, si $\hat{y}_0$ est loin de la valeur exacte, une valeur trop faible de P_0 entraîne une convergence très lente du processus.

2. IDENTIFICATION ET FILTRAGE NON LINEAIRE.

Supposons que la fonction f de (1.1) dépende d'un paramètre u (vecteur de dimensions p) :

$$(2.1) \quad y_{n+1} = f(y_n, n, u) + (\text{bruit dynamique})$$

l'observation étant toujours donnée par :

$$(2.2) \quad z_n = H_n y_n + (\text{erreur de mesure})$$

Considérons un nouveau vecteur d'état Y_n , de dimension $m + p$, défini par :

$$(2.3) \quad Y_n^i = y_n^i \quad i = 1, 2, \dots, m$$

$$(2.4) \quad Y_n^{m+\ell} = u_\ell \quad \ell = 1, 2, \dots, p$$

Ce nouveau vecteur satisfait à l'équation :

$$(2.5) \quad Y_{n+1} = F(Y_n, n) + (\text{bruit dynamique})$$

avec :

$$(2.6) \quad \left\{ \begin{array}{l} F_i(Y_n, n) = f_i(y_n, n, u) \end{array} \right. \quad i = 1, 2, \dots, m$$

$$(2.7) \quad \left\{ \begin{array}{l} F_{m+\ell}(Y_n, n) = Y_n^{m+\ell} \end{array} \right. \quad \ell = 1, 2, \dots, p$$

Les équations (2.7) entraînent bien que Y_n^ℓ , $\ell = m+1, \dots, m+p$,

(c'est-à-dire u) ne dépend pas de n . On peut encore définir l'observation par :

$$(2.8) \quad z_n = \mathcal{H}_n Y_n + (\text{erreur de mesure})$$

la matrice $\mathcal{H}_n$ dérivant de H_n comme suit :

$$(2.9) \quad \mathcal{H}_n = \left(\underbrace{H_n}_m \mid \underbrace{0}_p \right) \Bigg\}^q$$

On est ainsi ramené à un problème du type précédent, c'est-à-dire filter le système suivant :

$$\left\{ \begin{array}{l} Y_{n+1} = F(Y_n, n) + (\text{bruit dynamique}) \\ z_n = \mathcal{H}_n Y_n + (\text{erreur de mesure}) \end{array} \right.$$

L'estimation du paramètre u est donnée par les p dernières composantes du vecteur d'état y_n .

3. APPLICATIONS NUMERIQUES

3.1

Le problème continu.

Nous allons étudier un problème analogue à ceux de la première partie, mais avec un coefficient distribué cette fois. Plus précisément, considérons le système suivant :

- l'état vérifie :

$$(I) \left\{ \begin{array}{l} \frac{\partial y}{\partial t} - \frac{\partial}{\partial x} \left(u(x) \frac{\partial y}{\partial x} \right) = f \quad \text{sur } Q = (0,1) \times (0,T) \\ y(0,t) = y(1,t) = 0 \\ y(x,0) = y_0(x) \end{array} \right.$$

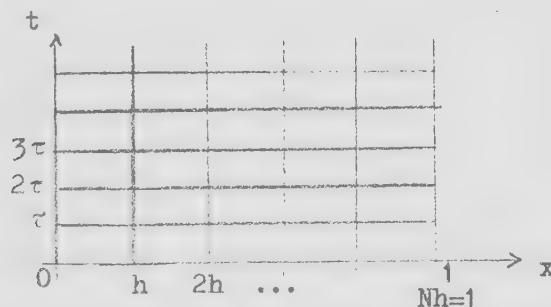
- l'observation z est donnée par :

$$(3.1) \quad z = \mathcal{C}y + \eta(t)$$

où $\mathcal{C}$ est une application linéaire continue de l'espace des états dans l'espace des observations et η un bruit blanc gaussien défini comme précédemment.

3.2 Le problème discrétisé.

On discrétise l'équation aux dérivées partielles (I) suivant un schéma explicite. Cela nous permet d'obtenir facilement une équation du type (1.1), assortie cependant d'une condition de stabilité qui nous oblige à prendre un pas de temps très faible. Considérons sur Q le réseau suivant :



Nous noterons :

y_n^i = valeur moyenne de y sur le pavé $(ih - \frac{3h}{2}, ih - \frac{h}{2}) \times (n\tau - \frac{3\tau}{2}, n\tau - \frac{\tau}{2})$

u_i = valeur moyenne de u sur l'intervalle $(ih - h, ih)$

f_n^i = valeur moyenne de f sur le pavé

$$(ih - \frac{3h}{2}, ih - \frac{h}{2}) \times (n\tau - \frac{3\tau}{2}, n\tau - \frac{\tau}{2})$$

L'équation (I) discrétisée s'écrit alors :

$$\frac{y_{n+1}^i - y_n^i}{\tau} - \frac{1}{h} \left[u_i \left(\frac{y_n^{i+1} - y_n^i}{h} \right) - u_{i-1} \left(\frac{y_n^i - y_n^{i-1}}{h} \right) \right] = f_n^i$$

$$i = 2, 3, \dots, N \quad ; \quad n = 1, 2, \dots$$

soit :

$$(3.2) \quad y_{n+1}^i = y_n^i + \frac{\tau}{h^2} \left[u_i y_n^{i+1} - (u_i + u_{i-1}) y_n^i + u_{i-1} y_n^{i-1} \right] + \tau f_n^i$$

$$i = 2, 3, \dots, N \quad ; \quad n = 1, 2, \dots$$

$$\text{avec } y_n^1 = y_n^{N+1} = 0 \quad n = 1, 2, \dots$$

$$(3.3) \quad y_1^i = \text{valeur moyenne de } y_0 \text{ sur l'intervalle } (ih - \frac{3h}{2}, ih - \frac{h}{2})$$

$$(3.3) \quad \text{Condition de stabilité :} \quad \tau \leq \frac{h^2}{2 \sup_{x \in (0,1)} u(x)}$$

L'observation devient :

$$z_n = H_n y_n + \xi_n$$

où ξ_n désigne une variable aléatoire gaussienne centrée d'écart-type σ . On se limitera à des observations ponctuelles, c'est-à-dire qu'on prendra pour H_n une matrice du type :

$$H_n = (0, 0, \dots, 1, 0, \dots, 0)_{j_n}$$

l'indice j_n dépendant seul de n . Le vecteur z_n est donc en fait réduit à un nombre réel.

3.3. Résultats numériques.

3.3.1. Le coefficient u est constant.

Nous avons programmé l'algorithme avec les données suivantes :

$$f(x,t) = \sin \pi x ((b - u_e \pi^2) e^{-bt} + a u_e \pi^2)$$

$$y_0(x) = (a-1) \sin \pi x .$$

la solution de (I) est , pour $u = u_e$:

$$y(x,t) = (a - e^{-bt}) \sin \pi x$$

Valeurs des différentes constantes :

$$a = 1.5 ; \quad b = 1 ; \quad u_e = 1.8$$

$$\text{pas en espace : } h = 1/10$$

$$\text{pas en temps : } \tau = 1/500$$

On observe la 6ème composante du vecteur d'état, pour tout n , c'est-à-dire $y(0.5, t)$.

Point de départ de l'algorithme :

- valeur initiale de l'estimateur : 1
- matrice P_0 : tous les éléments égaux à 1.

Les résultats sont rassemblés sur les figures suivantes :

Figure 1 : On a représenté l'estimateur $\hat{u}_n$ en fonction de n pour de faibles valeurs de n (sans bruit). On observe la convergence de $\hat{u}_n$ vers une valeur voisine de 1.8.

Figure 2 : C'est la même courbe pour les grandes valeurs de n . Cette figure met en évidence la stabilité de l'algorithme :

$$\hat{u}_n \xrightarrow{n \rightarrow \infty} \hat{u} = 1.815$$

Figure 3 : Nous considérons cette fois le système avec un bruit de valeur relative 1 %. Il y a deux courbes sur cette figure :

- (i) $1 \leq n \leq 110$: on représente comme précédemment $\hat{u}_n$ en fonction de n .

(ii) $110 \leq n$: des oscillations (très rapides) apparaissent sur la courbe précédente ; on représente alors la courbe moyenne :

$$\hat{u}_{n \text{ moyen}} = \frac{\hat{u}_{101} + \hat{u}_{102} + \dots + \hat{u}_n}{(n - 100)} \quad n \geq 100$$

Figure 4 : Ce sont les mêmes courbes que précédemment, pour un bruit de valeur relative 5 %. Pour des bruits plus importants, l'algorithme devient instable.

Temps de calcul : 10_s pour 500 itérations sur I.B.M 360/91 .

3.3.2. Le coefficient u est distribué : $u(x) = 1 + x$.

Si nous prenons pour second membre et condition initiale :

$$\begin{aligned} f(x,t) &= (\sin \pi x (b - (1+x)\pi^2) + \pi \cos \pi x) e^{-bt} \\ &+ a\pi^2(1+x) \sin \pi x - a\pi \cos \pi x \\ y_0(x) &= (a-1) \sin \pi x \end{aligned}$$

la solution, pour $u(x) = 1 + x$, de (I) s'écrit :

$$y(x,t) = \sin \pi x (a - e^{-bt})$$

Les constantes ont les mêmes valeurs que précédemment. On observe toujours une composante du vecteur d'état, mais on change de composante à chaque pas :

$$\begin{aligned} - \text{ pour } n \leq 9 & \quad j_n = n + 1 \\ - \text{ pour } n > 9 & \quad j_n = n[10] + 2 \end{aligned}$$

([] signifie modulo)

Point de départ de l'algorithme :

- valeur initiale de l'estimateur : $u_0(x) = 1$.
- matrice P_0 : tous les éléments sont pris égaux à 10 .

On a représenté sur les figures 5 et 6 les courbes $\hat{u}_n(x)$ pour

différents n .

Dans ce cas également, l'algorithme devient instable pour des bruits trop importants (σ_r supérieur à quelques pour cent).

Temps de calcul : 25 s. pour 500 itérations (sur I.B.M. 360/91).

On constate sur ces courbes que l'estimation est difficile au voisinage du point où la dérivée $\frac{\partial y}{\partial x}$ s'annule. Pour remédier à cela, on utilise un procédé issu du complément (voir 2.2) : on maintient égales les deux composantes de $\hat{u}_n$ encadrant le point précédent. Les figures 7 et 8 représentent les courbes ainsi obtenues. L'amélioration est très nette lorsqu'il n'y a pas de bruit ; elle est pratiquement inexistante par contre en présence de bruit.

En conclusion, la méthode des moindres carrés réursive est plus rapide que la méthode bayésienne et permet d'identifier des coefficients quelconques. Toutefois l'algorithme devient instable lorsque le bruit est trop important ; on n'a alors plus aucun résultat. De plus, on ne possède aucun théorème précis sur les estimateurs utilisés (sur la convergence en particulier).

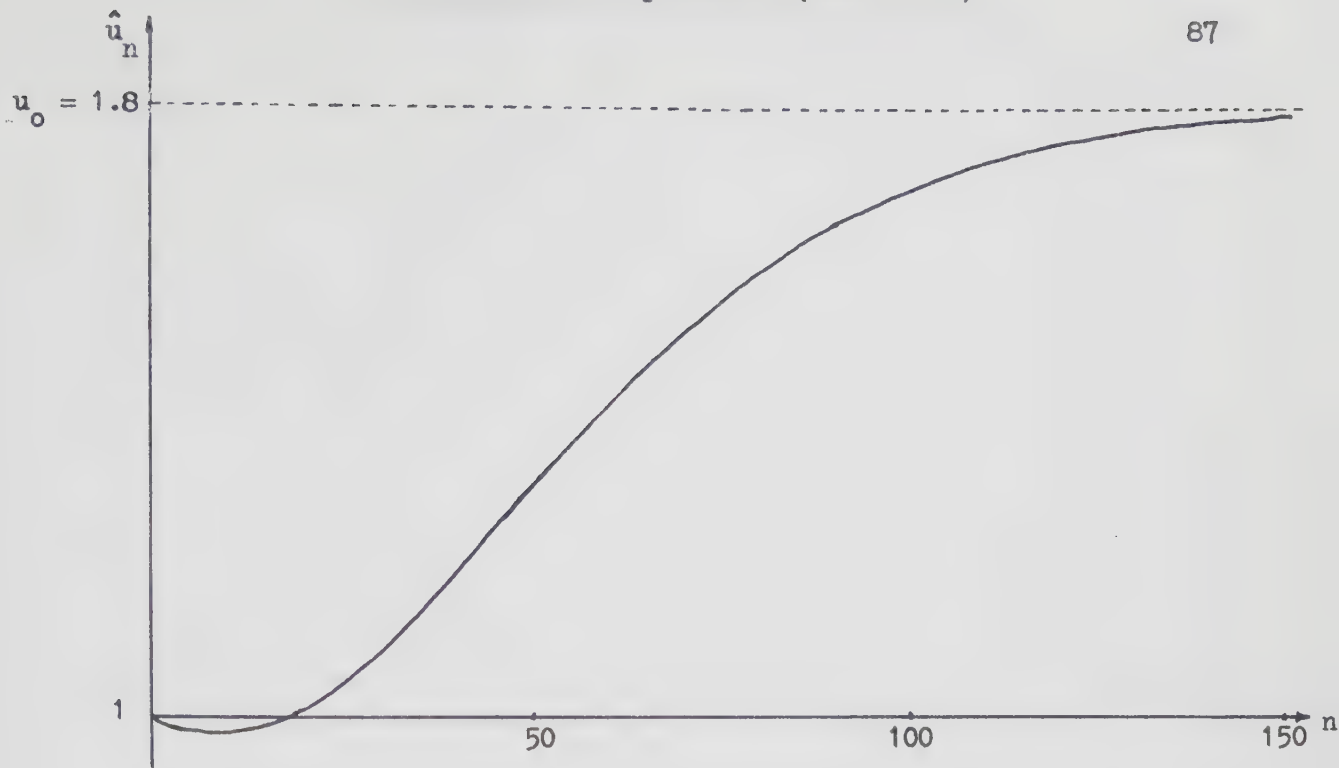


Figure 1 : $\hat{u}_n$ pour n petit

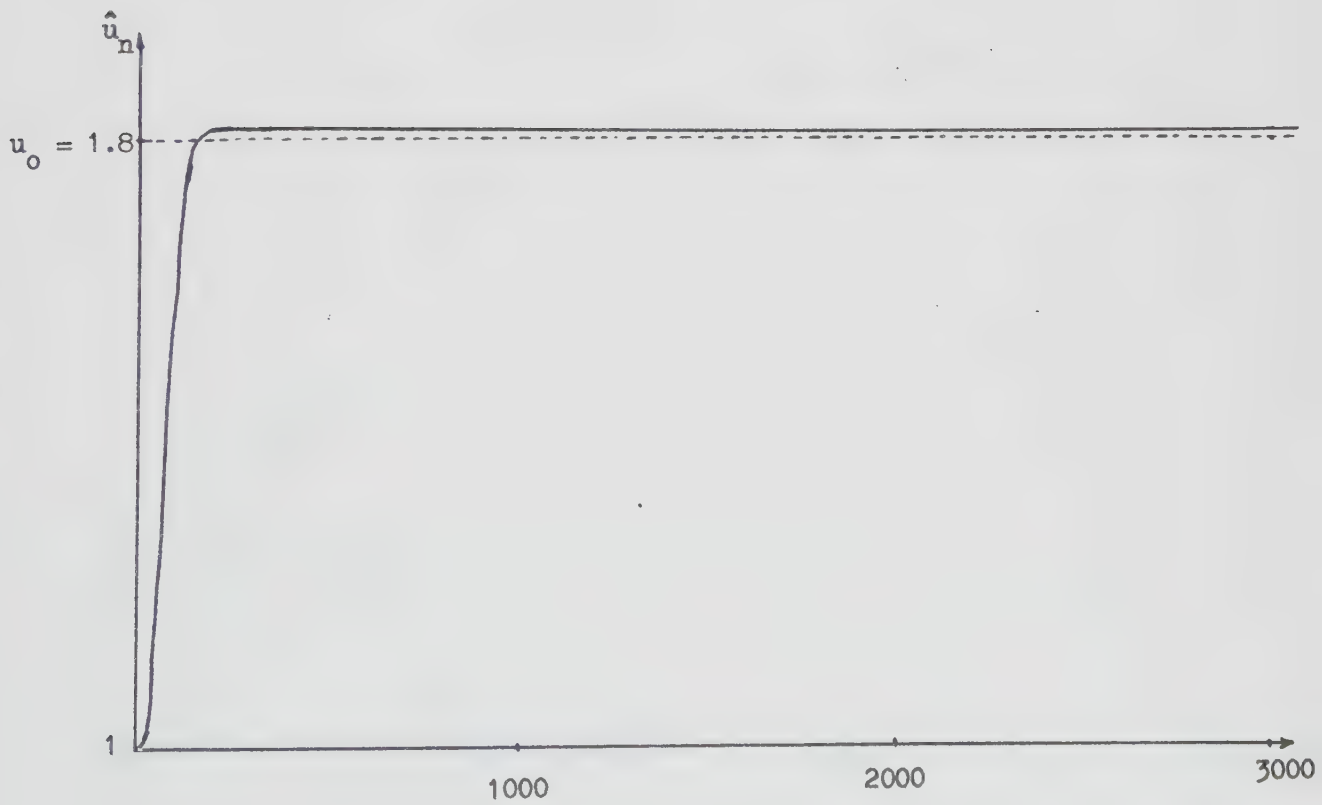


Figure 2 : $\hat{u}_n$ pour n grand

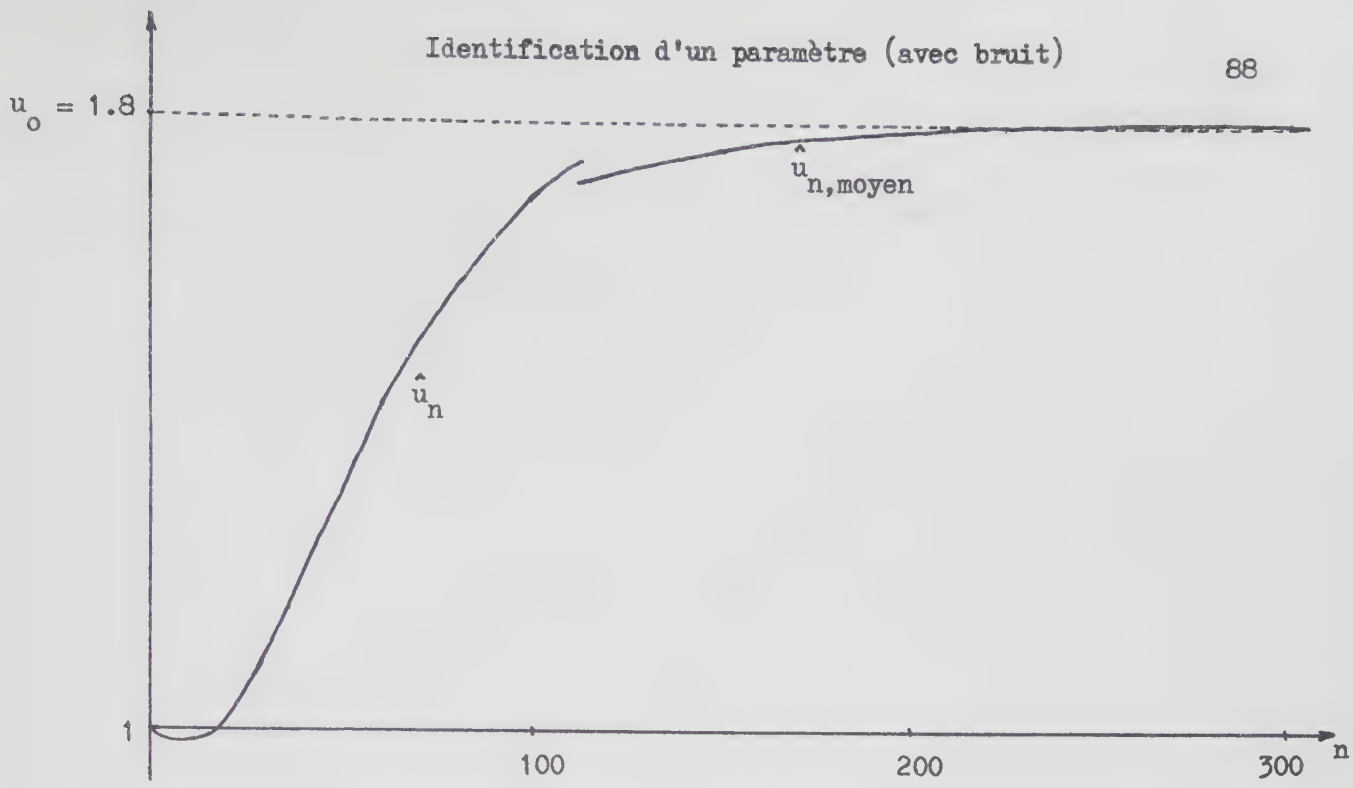


Figure 3 : $\sigma_r = 1 \%$

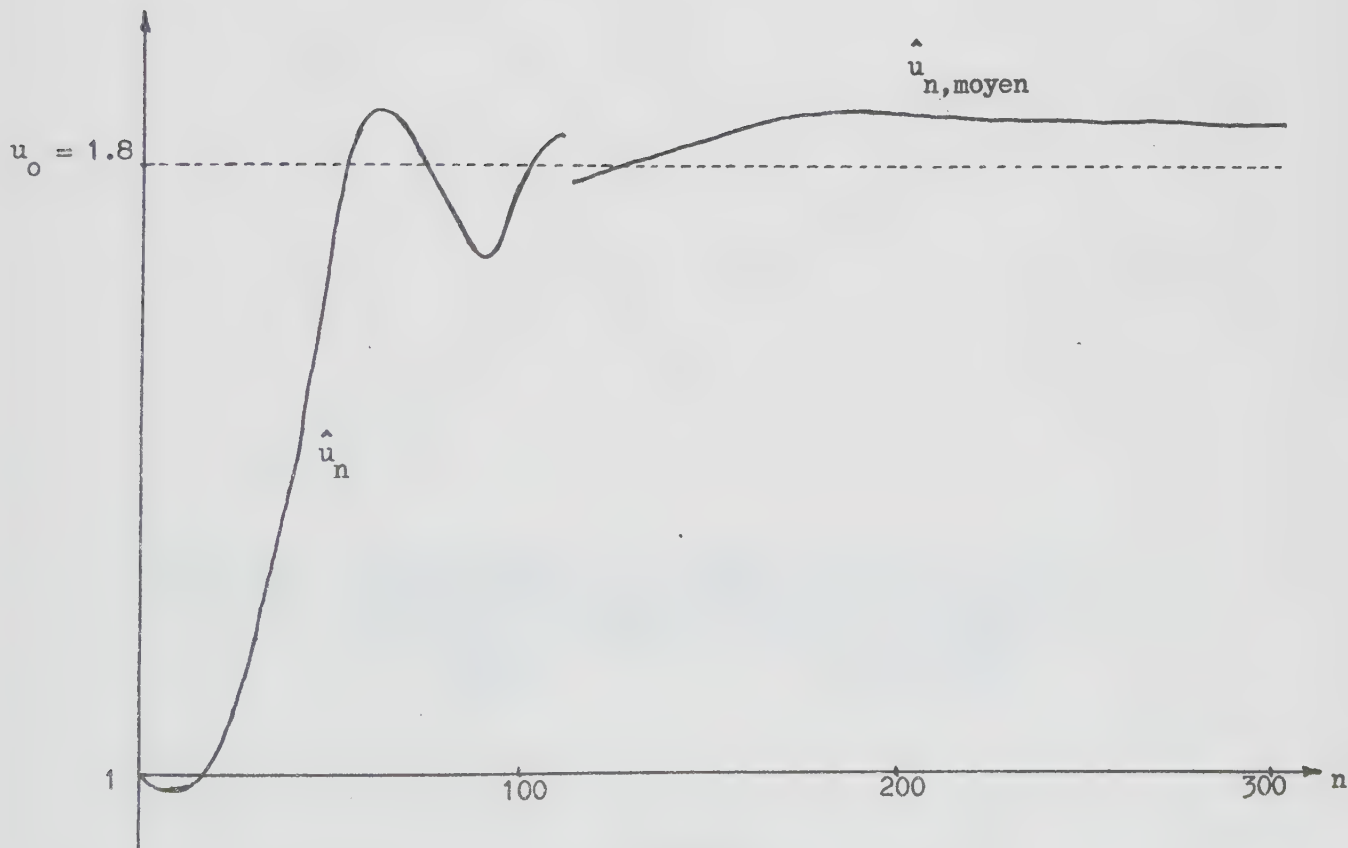


Figure 4 : $\sigma_r = 5 \%$

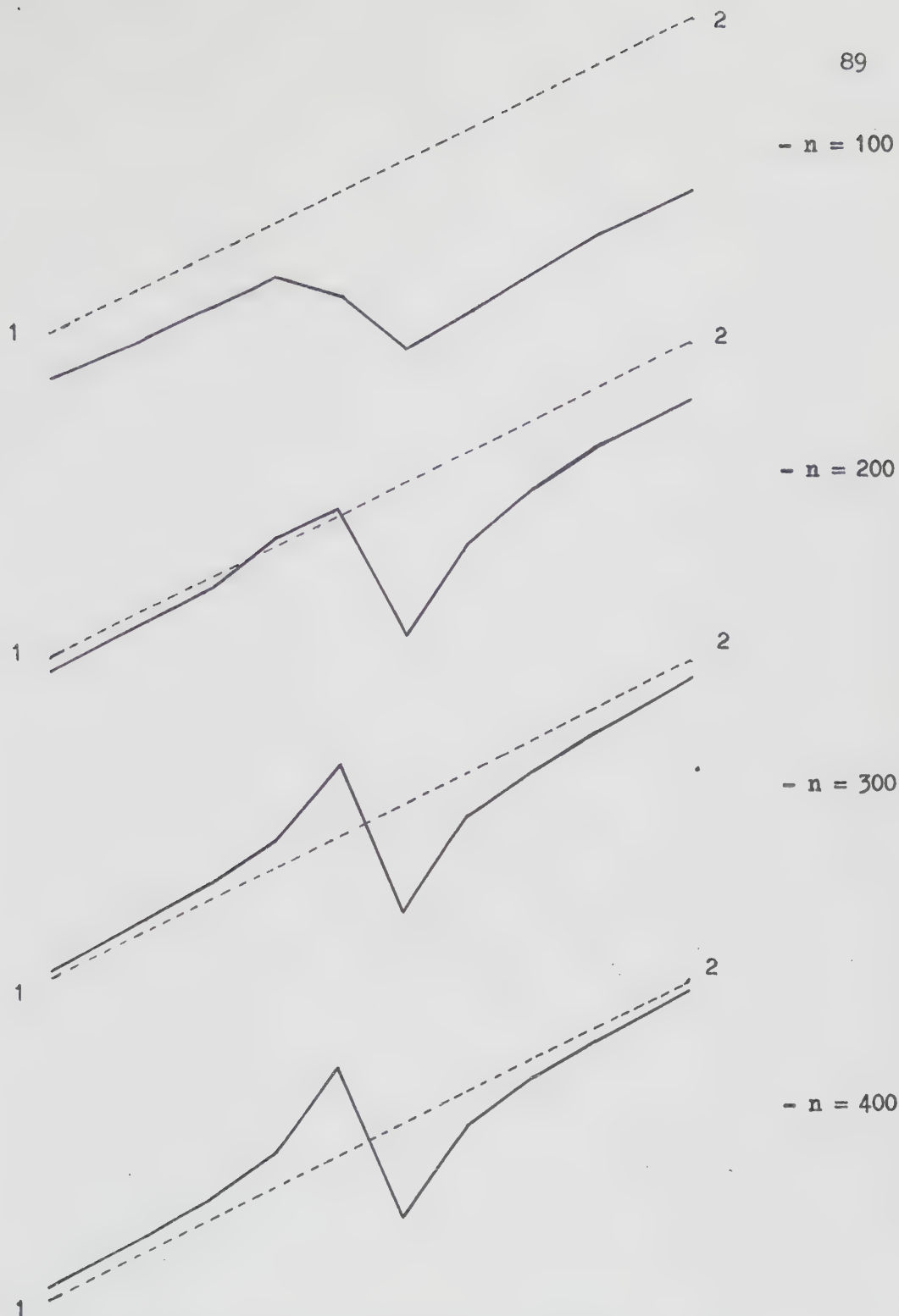


Figure 5 : Identification d'un paramètre distribué : $u_0(x) = 1 + x$
 (sans bruit). Estimation $\hat{u}_n(x)$ pour différents n .
 $u_0(x)$ — $\hat{u}_n(x)$



Figure 6 : Identification d'un paramètre destiné : $u_0(x) = 1 + x$
 (avec bruit : $\sigma_r = 1\%$). Estimation $\hat{u}_n$ pour différents n .
 $u_0(x)$ — $\hat{u}_n(x)$



Figure 7 : Identification d'un paramètre distribué : $u_0(x) = 1 + x$.
Estimateur adapté, sans bruit.

..... $u_0(x)$ _____ $\hat{u}_n(x)$

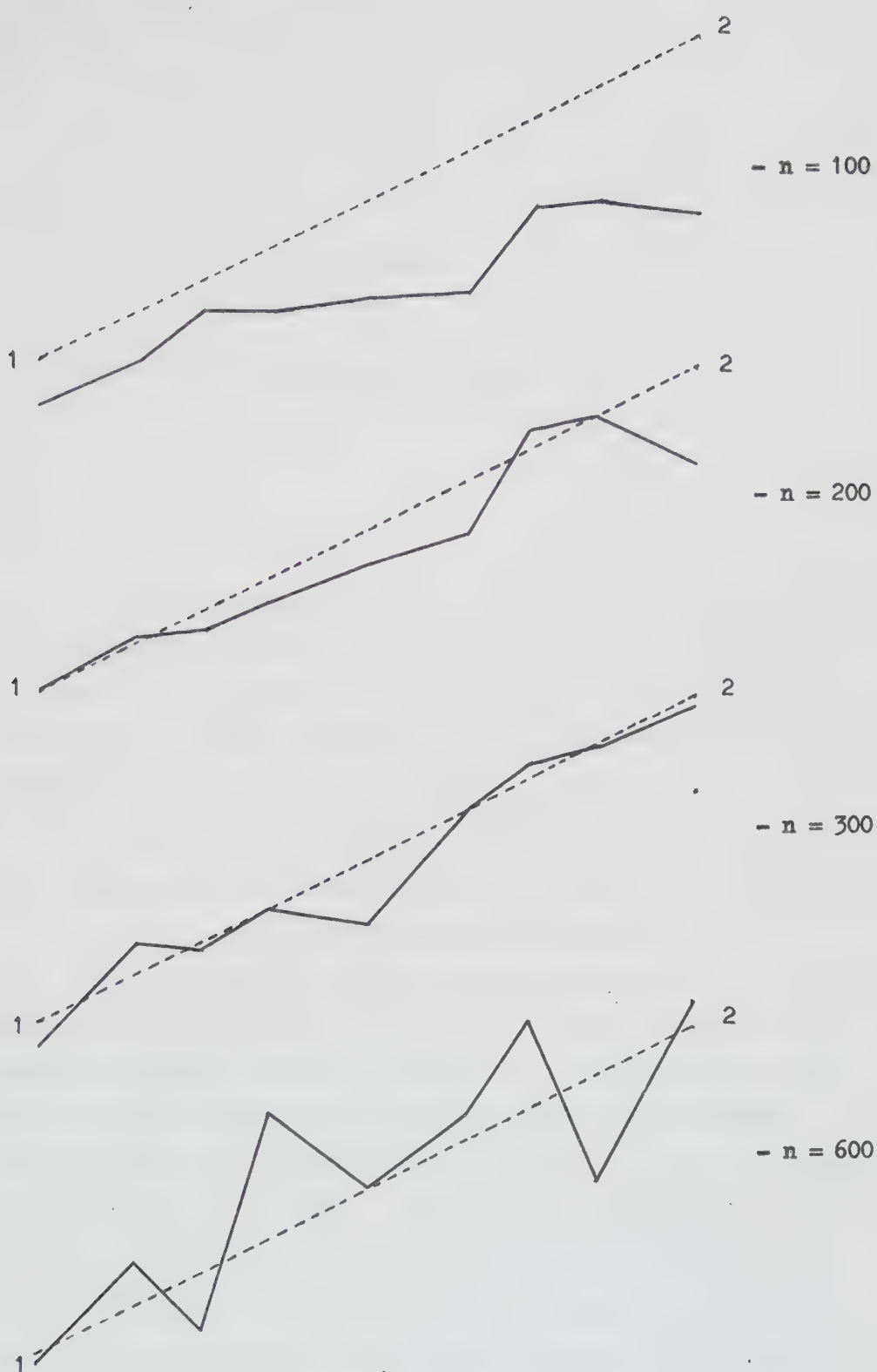


Figure 8 : Identification d'un paramètre distribué : $u_0(x) = 1 + x$.

Estimateur adapté, avec bruit : $\sigma_r = 1\%$

..... $u_0(x)$ — $\hat{u}_n(x)$

COMPLEMENT

UNE METHODE DIRECTE

Pour compléter l'étude des différentes techniques d'identification des systèmes gouvernés par des équations aux dérivées partielles, nous terminerons par les méthodes directes, méthodes simples mais applicables uniquement sous les conditions restrictives suivantes :

- (i) On suppose que l'on observe directement l'état du système.
- (ii) Cette observation n'est perturbée par aucun bruit. (En fait, on peut seulement tolérer un bruit très faible).

Les techniques précédentes, bien sûr applicables, sont lourdes et trop peu précises dans ce cas. Il semble préférable d'utiliser une méthode directe qui consiste à se ramener à la résolution d'un système d'équations linéaires. Signalons une première méthode proposée par COLLINS-KHATRI dans [5] . Des essais numériques de cette technique ont donné de bons résultats. Nous avons traité, en particulier, l'exemple 3.3.2 de la deuxième partie obtenant $u(x)$ avec une excellente précision (erreur inférieure à 10^{-6}) pour un temps de calcul de l'ordre de 2 s (I.B.M. 360/91) à partir d'un réseau de discrétisation (20×20). Nous proposons, dans ce complément, une autre méthode développée à partir de l'équation de la diffusion.

1. POSITION DU PROBLEME

Considérons une équation aux dérivées partielles représentée sous la forme :

$$P(u) \cdot y = f$$

Le problème habituel consiste à trouver y connaissant u , f et les conditions aux limites sur y . Considérons maintenant l'équation par rapport à u , que nous supposons réparti. Il s'agit de déterminer u connaissant y et f . Il n'y a pas en général de condition aux limites sur u . Mis ainsi hors du cadre classique, nous n'avons pas de théorèmes généraux d'existence et d'unicité. Ces difficultés se retrouvent au niveau de l'approximation : le problème discrétisé se ramène à un système d'équations linéaires comportant plus d'inconnues que d'équations. Nous limiterons notre étude au problème de la diffusion stationnaire isotrope, auquel cas des hypothèses de régularité sur u et y permettent d'obtenir l'unicité (théorique et numérique).

2. EQUATION DE LA DIFFUSION A UNE DIMENSION

2.1 Le problème continu

L'équation de la diffusion, avec des conditions aux limites homogènes, s'écrit :

$$(2.1) \quad \begin{cases} \frac{d}{dx}(u(x) \frac{dy}{dx}) = f & \text{sur } \Omega =]0,1[\\ (I) \end{cases}$$

$$(2.2) \quad \begin{cases} y(0) = y(1) = 0 \end{cases}$$

On sait que, si f est dans $L^2(\Omega)$ et u dans $C^1(\overline{\Omega})$, alors y existe et est unique dans $H^2(\Omega)$. Soit y_0 la solution de (I) pour $u = u_0(x)$. Il s'agit maintenant, à partir de y et f , de retrouver u_0 . Tel que nous avons posé le problème, il possède

au moins une solution dans $C^1(\bar{\Omega})$. Montrons qu'elle est unique. Pour cela, intégrons l'équation (2.1) ; on obtient, si $F(x)$ désigne une primitive de $f(x)$:

$$(2.3) \quad u(x) \frac{\partial y}{\partial x} = F + C$$

où C est une constante arbitraire. Nous savons que :

$$y \in H^2(\Omega) \subset C^1(\bar{\Omega})$$

Donc il existe x_0 de $]0,1[$ tel que $\frac{\partial y}{\partial x}(x_0) = 0$.

Alors, si $x \rightarrow x_0$

$$u(x) \rightarrow u(x_0) < \infty \quad \text{car } u \in C^1(\bar{\Omega})$$

$$\frac{\partial y}{\partial x}(x) \rightarrow 0 \quad \text{puisque } y \in C^1(\bar{\Omega})$$

$F(x) \rightarrow F(x_0)$, F est continue, car c'est une primitive.

il faut alors que $C = -F(x_0)$, sinon $u(x) \rightarrow \infty$ quand $x \rightarrow x_0$.

Nous avons ainsi l'unicité de u_0 dans $C^1(\bar{\Omega})$.

2.2 Le problème discrétisé

Nous avons étudié l'équation (I) avec les fonctions suivantes :

$$(2.4) \quad \begin{cases} u_0(x) = 1 + x \\ f(x) = \pi \cos \pi x - \pi^2(1+x) \sin \pi x \\ \text{soit } y_0(x) = \pi \sin \pi x \end{cases}$$

Considérons un schéma d'approximation classique aux différences finies.

Soit $(x_i, i = 1, \dots, N+1)$ un découpage de $(0,1)$:

$$\begin{array}{ccccccc} 0 & h & 2h & & (N-1)h & 1 & = Nh \\ \hline x_1 & x_2 & & & x_N & x_{N+1} & \end{array}$$

Posons : y_i = valeur moyenne de y sur l'intervalle $(x_i - \frac{h}{2}, x_i + \frac{h}{2})$

u_i = valeur moyenne de u sur l'intervalle (x_i, x_{i+1})

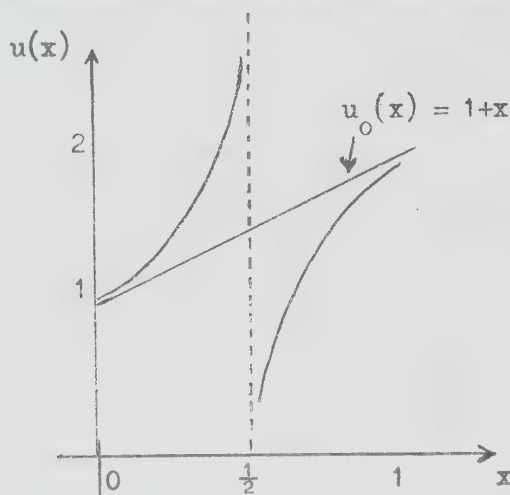
f_i = valeur moyenne de f sur l'intervalle $(x_i - \frac{h}{2}, x_i + \frac{h}{2})$

L'équation (I) discrétisée s'écrit, par rapport aux u :

$$(2.5) \quad u_{i-1}(y_{i-1} - y_i) + u_i(y_{i+1} - y_i) = h^2 f_i \quad i = 2, \dots, N$$

On a ainsi $N-1$ équations pour N inconnues, $u_1, u_2, \dots, u_N$. Il faut donc en ajouter une autre. Dans le paragraphe précédent (2.1), on a vu que l'unicité résultait du fait que $u(x)$ ne peut tendre vers l'infini lorsque $\frac{\partial y}{\partial x}$ s'annule. Nous allons utiliser cette propriété. Dans notre exemple $\frac{\partial y}{\partial x}$ change de signe en s'annulant au point $x = \frac{1}{2}$. Si

la constante C est mal choisie, nous obtenons une courbe $u(x)$ à l'allure suivante :



Pour éviter ces branches infinies, nous fixons égales les deux composantes de u situées de part et d'autre de $\frac{1}{2}$, c'est-à-dire :

$$(2.6) \quad u_5 = u_6$$

Nous pouvons maintenant résoudre le système formé de (2.5) et (2.6). La forme particulière de (2.6) nous permet de le faire de proche en proche, à partir de u_5 et u_6 .

2.3 Résultats numériques.

Avec un pas de $1/20$ et les fonctions données en (2.4), nous avons obtenu les résultats suivants :

x	0.025	0.075	0.125	0.175	0.225	0.275	0.325	0.375	0.425	0.5
$u_0(x)$	1.025	1.075	1.125	1.175	1.225	1.275	1.325	1.375	1.425	1.5
u calculé	1.028	1.079	1.129	1.179	1.230	1.280	1.331	1.383	1.436	1.503
x	0.575	0.625	0.675	0.725	0.775	0.825	0.875	0.925	0.975	
$u_0(x)$	1.575	1.625	1.675	1.725	1.775	1.825	1.875	1.925	1.975	
u calculé	1.570	1.623	1.675	1.726	1.776	1.827	1.877	1.927	1.978	

Avec un bruit sur y de valeur relative $1/1000$:

$u_0(x)$	1.025	1.075	1.125	1.175	1.225	1.275	1.325	1.375	1.425	1.5
u calculé	1.007	1.054	1.108	1.167	1.205	1.227	1.292	1.320	1.350	1.319
$u_0(x)$	1.575	1.625	1.675	1.725	1.775	1.825	1.875	1.925	1.975	
u calculé	1.534	1.679	1.720	1.750	1.830	1.851	1.902	1.948	1.996	

Pour des bruits plus importants, on n'obtient plus de résultats valables.

3. EQUATION DE LA DIFFUSION EN DIMENSION DEUX

3.1 Le problème continu

En dimension deux, l'équation de la diffusion s'écrit :

$$\text{II} \quad \left\{ \begin{array}{l} \frac{\partial}{\partial x_1} (u(x_1, x_2) \frac{\partial y}{\partial x_1}) + \frac{\partial}{\partial x_2} (u(x_1, x_2) \frac{\partial y}{\partial x_2}) = f \quad \text{sur } \Omega =]0,1[\times]0,1[\\ y(0, x_2) = y(1, x_2) = y(x_1, 0) = y(x_1, 1) = 0 \end{array} \right.$$

Si nous prenons :

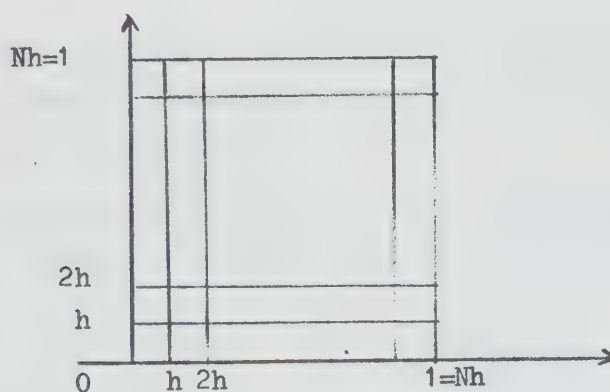
$$(3.1) \quad \left\{ \begin{array}{l} u_0(x_1, x_2) = 1 + x_1 + x_2 \\ f(x_1, x_2) = -2\pi^2(1+x_1+x_2)\sin\pi x_1 \sin\pi x_2 + \pi \sin\pi x_2 \cos\pi x_1 + \pi \sin\pi x_1 \cos\pi x_2 \\ \text{on obtient } y_0(x_1, x_2) = \sin\pi x_1 \sin\pi x_2 \end{array} \right.$$

Connaissant f et y_0 , il s'agit comme précédemment de retrouver u_0 .

Ainsi posé, il y a évidemment existence. L'unicité, dans $C^\infty(\overline{\Omega})$ cette fois, a été démontrée par CHAVENT (cf. [3]). Sans entrer dans les détails, la démonstration repose encore sur le fait que, si y présente un extremum sur Ω , la régularité de u ($u \in C^\infty(\overline{\Omega})$) entraîne l'unicité.

3.2 Discrétisation

Sur le carré $(0,1) \times (0,1)$, on dispose un réseau régulier :



On pose : $x_{ij} = ((i-1)h, (j-1)h)$ $i, j = 1, 2, \dots, N+1$

$$y_{ij} = y(x_{ij}) \quad "$$

$$f_{ij} = f(x_{ij}) \quad "$$

$$u_{ij} = u((i-1)h + \frac{h}{2}, (j-1)h + \frac{h}{2}) \quad i, j = 1, \dots, N.$$

L'équation (II) devient :

$$(3.2) \quad \frac{1}{2h^2} \left[(u_{i,j} + u_{i,j-1})(y_{i+1,j} - y_{i,j}) - (u_{i-1,j} + u_{i-1,j-1})(y_{i,j} - y_{i-1,j}) \right. \\ \left. + (u_{i,j} + u_{i-1,j})(y_{i,j+1} - y_{i,j}) - (u_{i,j-1} + u_{i-1,j-1})(y_{i,j} - y_{i,j-1}) \right] = f_{ij} \\ i, j = 2, 3, \dots, N.$$

soit, ordonné par rapport aux u_{ij} :

$$u_{i-1,j-1}(y_{i,j-1} - 2y_{i,j} + y_{i-1,j}) + u_{i,j-1}(y_{i+1,j} - 2y_{i,j} + y_{i,j-1}) \\ + u_{i-1,j}(y_{i,j+1} - 2y_{i,j} + y_{i-1,j}) + u_{i,j}(y_{i+1,j} - 2y_{i,j} + y_{i,j+1}) = 2h^2 f_{ij} \\ i, j = 2, \dots, N$$

On est ainsi ramené à un système de $(N-1)^2$ équations pour N^2 inconnues; $2N-1$ équations supplémentaires sont donc nécessaires, Nous avons choisi les relations suivantes :

(i) au voisinage de l'extremum de y_0 , on maintient égaux les quatre u_{ij} voisins, c'est-à-dire, si y est extrémale en x_{io,j_0} :

$$(3.3) \quad \begin{array}{ccc} u_{io-1,j_0} & \begin{array}{|c|c|c|} \hline & & \\ \hline \bullet & & \bullet \\ \hline \end{array} & u_{io,j_0} \\ u_{io-1,j_0-1} & \begin{array}{|c|c|c|} \hline & & \\ \hline \bullet & \bullet & \bullet \\ \hline \end{array} & u_{io,j_0-1} \\ & \begin{array}{c} \uparrow \\ x_{io,j_0} \end{array} & \end{array} \quad u_{io-1,j_0} = u_{io-1,j_0-1} = u_{io,j_0-1} = u_{io,j_0}$$

$$(3.4) \quad \text{dans notre exemple, cela donne :} \\ u_{5,6} = u_{5,5} = u_{6,5} = u_{6,6}$$

(ii) Le long des courbes où $\frac{\partial y}{\partial x_1}$ et $\frac{\partial y}{\partial x_2}$ s'annulent, on impose aux u_{ij} d'être égaux par paires. Dans notre cas $\frac{\partial y}{\partial x_2}$ s'annule sur la droite $x_1 = \frac{1}{2}$, $\frac{\partial y}{\partial x_1}$ sur $x_2 = \frac{1}{2}$; nous avons alors les relations :

$$(3.5) \quad u_{5,j} = u_{6,j} \quad j = 1, \dots, N$$

$$(3.6) \quad u_{i,5} = u_{i,6} \quad i = 1, \dots, N$$

La partie (i) nous apporte trois équations, (ii) en apporte $2(N-2)$. Nous avons alors les $2N-1$ relations nécessaires. L'intérêt de ces conditions est de permettre la résolution du système ainsi formé de proche en proche ; nous évitons ainsi d'utiliser les méthodes itératives habituelles et les problèmes de conditionnement qui s'y rattachent.

3.3

Résultats numériques

On a programmé l'exemple ci-dessus avec un pas de $1/20$ en x_1 , et $1/20$ en x_2 . Les résultats sont rassemblés sur les figures 1 et 2:

Figure 1 : On a représenté trois coupes de l'erreur relative Δr

$$\Delta r = \frac{|u_o(x_1, x_2) - u_{\text{calculé}}|}{u_o(x_1, x_2)}$$

Figure 2 : Nous supposons que l'observation est perturbée par un bruit gaussien d'écart-type relatif $1/1000$; nous avons tracé des coupes de la courbe $u_{\text{calculé}}$

Temps de calcul : 6 s sur I B M 360/91 .

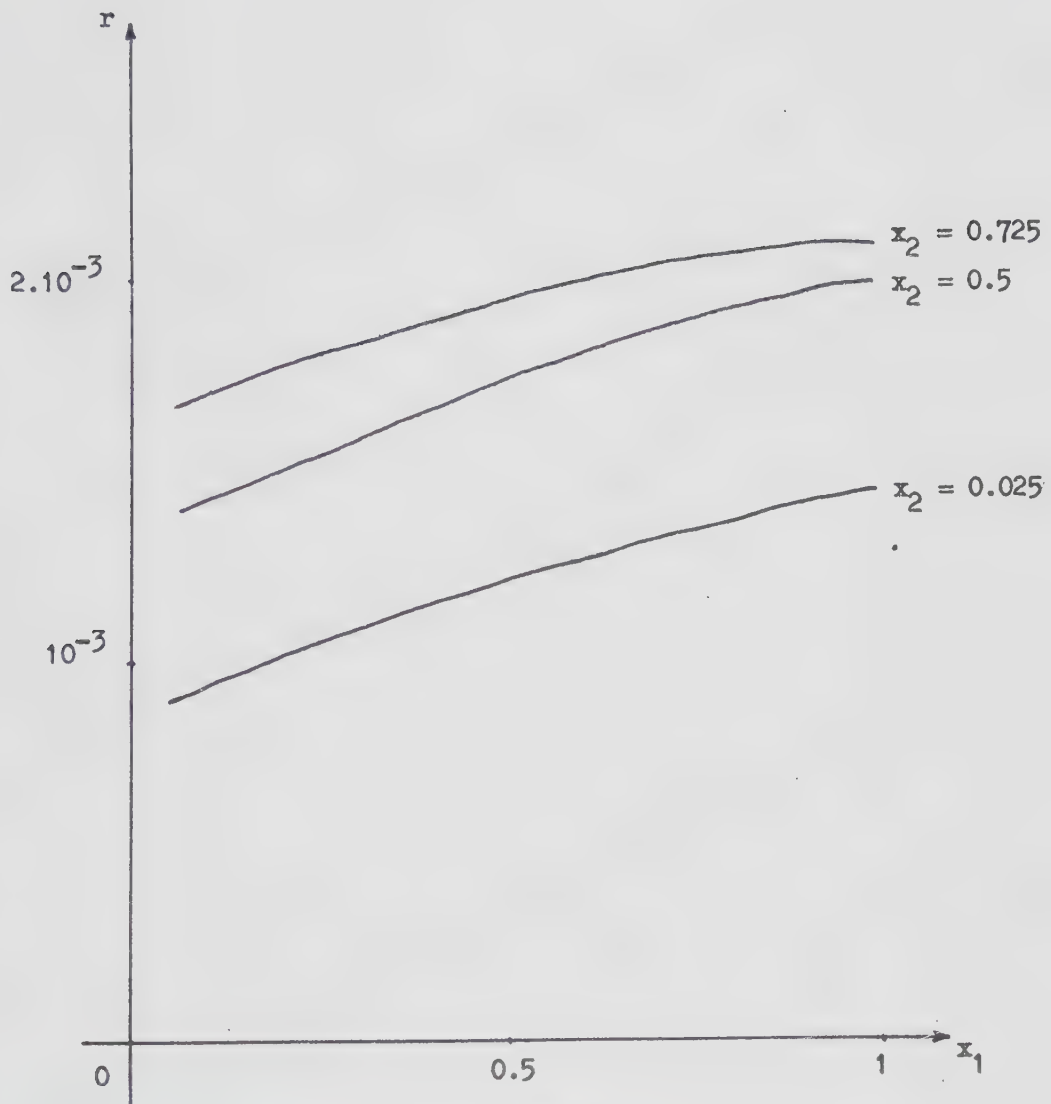


Figure 1

Identification d'un paramètre distribué à deux dimensions :

$$u_0(x_1, x_2) = 1 + x_1 + x_2$$

Coupe de l'erreur relative Δr pour différentes valeurs de x_2 .

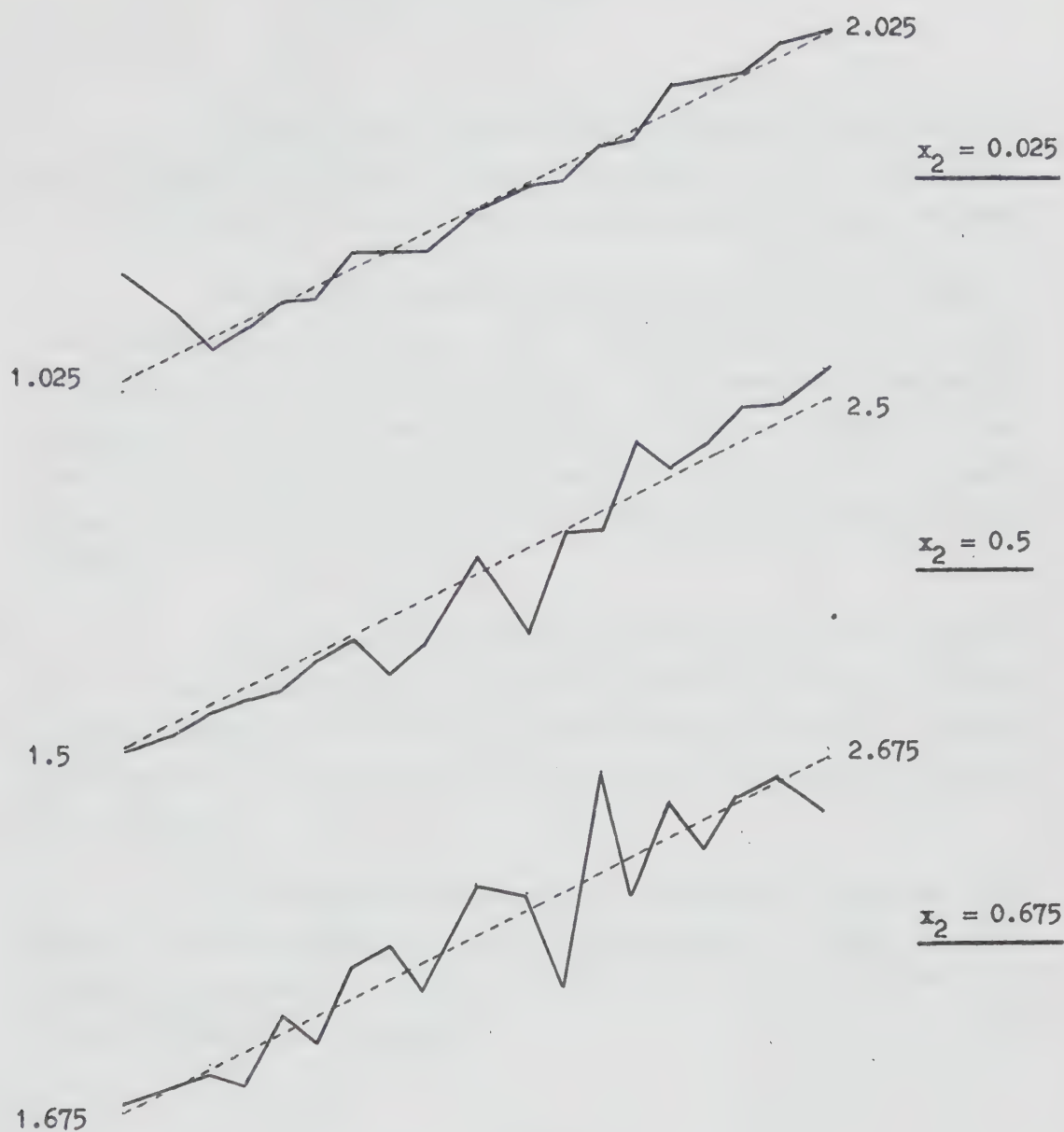


Figure 2

Identification d'un paramètre distribué à deux dimensions.

Coupe de $u_{\text{estimé}}$ pour différentes valeurs de x_2 .

..... $u_0(x_1, x_2) = 1 + x_1 + x_2$

———— $u_{\text{estimé}}(x_1, x_2)$

CONCLUSION G E N E R A L E

Il nous reste à comparer les différentes méthodes présentées ; on pourra utiliser, à cet effet, l'exemple (3.3.2. IIème partie) qui a été traité, sous des formes très voisines par les trois méthodes.

La méthode bayésienne exige le calcul de la loi a posteriori, ce qui limite son application à l'identification de deux ou trois paramètres constants, mais fournit, en contrepartie, beaucoup de renseignements (la variance, par exemple, qui donne une idée de la valeur des résultats). En plus de la stabilité numérique, elle est seule à donner des estimateurs qui convergent vers la valeur vraie des paramètres.

L'algorithme récursif des moindres carrés présente les caractéristiques inverses. Il est rapide et permet d'identifier des paramètres distribués. Par contre, il devient instable pour des bruits importants et ne donne aucun renseignement sur la qualité de l'estimation obtenue.

Les méthodes directes, très performantes (précision et rapidité nettement supérieures à celles des méthodes précédentes), présentent, hélas, des conditions d'applications très restrictives qui limitent beaucoup leur intérêt.

BIBLIOGRAPHIE

- [1] BENSOUSSAN (A), Sur l'identification et le filtrage de systèmes gouvernés par des équations aux dérivées partielles, thèse, Paris (1969).
- [2] CEA (J), Optimisation, cours de l'école d'été d'analyse numérique de l'E. D. F. (Juillet 1969).
- [3] CHAVENT (G), C.R. Acad. Sc. Paris, t.269, Série A, p 1135-1138; t.270, série A, p 25-28.
- [4] CHAVENT (G), Thèse à paraître.
- [5] COLLINS (P.L.) et KHATRI (H.C.), Identification of distributed parameter systems using finite differences, Journal of Basic Engineering (June 1969 p.239-245).
- [6] EMSELLEM (Y), Présentation d'une méthode nouvelle de résolution du "problème inverse" , laboratoire d'hydrologie mathématique de l'Ecole des Mines de Paris (5 Mai 1969).
- [7] LIONS (J.L.), Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles, Dunod, Paris (1968).
- [8] LIONS (J.L.) et MAGENES (E), Problèmes aux limites non homogènes et applications, volume 1, Dunod, Paris (1968).
- [9] NAHI (N.E.), Estimation theory and applications, J. Wiley (1969).
- [10] SRIDHAR (R.) et PEARSON (J.B.), Digital estimation of non linear processes.

DEUXIÈME THÈSE

Propositions données par la Faculté

" Equations de Carleman ."

Vu et approuvé

Paris, le ... 18 Juin 19...70

Le Doyen de la Faculté des Sciences
de Paris

Marc. ZAKHARSKI

Vu
et permis d'imprimer

Le Recteur de l'Académie de Paris

Robert WALLET

